**Knowledge Management Research & Practice (2010) 8,** 193–202
© 2010 Operational Research Society. All rights reserved 1477–8238/10
www.palgrave-journals.com/kmrp/

# Social semantic cloud of tags: semantic model for folksonomies

Hak-Lae Kim[1]
John Breslin[2]
Hong-Gee Kim[3] and
Jae-Hwa Choi[4]

*[1]Service Platform Group, Media Solution Center, Digital Media & Communications Business, Samsung Electronics Co., LTD., Suwon-city, Korea; [2]Electrical and Electronic Engineering, School of Engineering and Informatics, National University of Ireland, Galway, Ireland; [3]Biomedical Knowledge Engineering Lab, School of Dentistry, Seoul National University, Seoul, Korea; [4]Dankook University, College of Business, Choongnam, Korea*

**Correspondence: Hak-Lae Kim, Service Platform Group, Media Solution Center 416, Maetan-3Dong, Yeongtong-Gu, Suwon-Si Gyeonggi-Do, 443-742, Korea.**
**Tel: +82 10 3201 0714;**
**Fax: +82 31 279 8472;**
**E-mail: haklae.kim@gmail.com**

**Received: 17 December 2009**
**Accepted: 22 February 2010**

**Abstract**
A growing number of tagging applications have begun to provide users the ability to socialise their own keywords. Tagging, which assigns a set of keywords to resources, has become a powerful way for organising, browsing, and publicly sharing personal collections of resources on the Web. It is called folksonomies. These systems on current social websites, however, have deficiencies in defining tag's meaning, and are often blocked to users in order to reuse, share, and exchange the tags across heterogeneous websites. In this paper, we describe a semantic model for expressing folksonomies in social websites. This model, called Social Semantic Cloud of Tags, aims to provide a consistent format of representing folksonomies and some features in terms of tagging activities. We describe core concepts and relevant properties such as a popularity and usage of tags, along with deduced relationships between tags. We will discuss how this model helps to reduce drawbacks regarding tag sharing between users, applications, or folksonomies.
*Knowledge Management Research & Practice* (2010) **8,** 193–202.
doi:10.1057/kmrp.2010.10

**Keywords:** conceptual modeling; explicit knowledge; folksonomy; ontology; semantic web; tagging

## Introduction

The influence of the Web has been remarkable in the areas of computing infrastructure such as database, information retrieval, visualisation, etc. (Hendler *et al.*, 2008). Recently, many interesting aspects of the use of the Web have been widely popularised, including tagging and folksonomies, collaborative filtering, social bookmarking, social networking services, social search, mash-ups, and micro-blogging. The popular social websites provide their users common functions to maximise social interactions through various objects (e.g., a bookmark (Delicious[1]), a photo (Flickr[2]), a video clip (YouTube[3]), or a profile (Facebook[4]). Through these services people take part in the content creation and sharing, and they can make social connections between themselves. The self-motivated participation of people leads to new opportunities in the use of the Web. We, however, encounter some of the technical and social challenges due to its continuing growth and impact. Although the social software and technologies look very promising and potentially fit for the purpose of making social networks between people, yet a number of issues such as interoperability between independent platforms, data representation, and portability arise.

[1]http://delicious.com
[2]http://www.flickr.com
[3]http://www.youtube.com
[4]http://www.facebook.com

This paper focuses on *social tagging*. Tagging refers to an activity in which users assign a set of keywords to a resource. It is one of the common methods for participating on Social Web, and is also a central medium leading to social interaction (Quintarelli, 2005; Weinberger, 2005). A tag is a keyword that acts like a subject or category for the associated content. The number of social software applications offering tagging facilities is steadily increasing. The Pew Internet & American Life reported that (Rainie, 2007), '28% of Internet users have tagged or categorised content online such as photos, news stories or blog posts. On a routine day online, 7% of Internet users answered that they tag or categorise online content'. Gartner has also identified tagging as one of the seven core benefits that have come out of Web 2.0 technologies (Nadler, 2006).

In spite of a growing interest in terms of social tagging, inherent deficiencies include linguistic and grammatical variations as well as human typing errors. Furthermore, tagging practices that collect individual tagging activities are locked into a specific social website, meaning that a social website does not allow users from different platforms to access or share them (Kim *et al*., 2008a). As the number of folksonomic platforms increases, the complexity of folksonomies increases, while the interoperation of folksonomies between independent platforms decreases. To solve these problems, an alternative technology can be considered. Ontology-based Semantic Web technologies provide standards to interlink diverse folksonomic platforms, facilitating machine-readable metadata on web content.

In this paper, we propose a metadata model for representing folksonomies using Semantic Web technologies. This model is called Social Semantic Cloud of Tags (SCOTs), which represents folksonomies, including the tags, the resources that are being tagged, and the users that create these tags. We begin with giving an overview of our motivations for developing the SCOT ontology, followed by an introduction to conceptual foundations in this area.

## Motivation

In this section, we discuss a number of issues with respect to tagging and folksonomies. Shirky (2005) places emphasis on the folksonomy as an emergent pattern of users' collective intelligence, and Vander Wal (2005) claims that folksonomies can be harnessed to create a bottom-up and emergent view of the world. A tag offers a quick, simple, and easy way for organising information (Mathes, 2004; Merholz, 2004; Kroski, 2005), and allows for the exchange of their shared interests (Gruber, 2008). Over time, users may build up a rich collection of tags on a certain site or application.

The major drawbacks of current tagging systems are lack of semantics and keyword ambiguity (Mathes, 2004). A tag can be represented by a number of variations such as capitalisation (e.g., *Apple* and *apple*), singular *vs* plural (e.g., *blog* and *blogs*), or delimited words (e.g., *iPhone* and *i-phone*), while the relationships between tags are not expressed in an explicit manner. From a sharing point of view, current tagging systems provide restricted functionalities for users to share tagging data. The users registering on a certain site such as Delicious can share their tags and resources, and construct social networks using their tags. However, since these tagging practices are locked into the site, users cannot share, manipulate, and mix their data for the purpose of tag sharing (TagCommons, 2007a). This problem is related not only to the policies of the site, but also to describing tagging behaviours in a consistent way. Presently, most tagging systems do not provide a standardised format to share, exchange, and reuse tag data among users or communities. Although these systems offer Really Simple Syndication (RSS) and public Application Programming Interfaces (APIs) to publish their data, there is no formal conceptualisation to represent tagging data in a consistent way, and there exists no interoperability support for exchanging tagging data among different applications or people (TagCommons, 2007b).

## Tagging with semantics

There are a number of debates currently on the merits of folksonomies and traditional classifications. Shirky (2005) makes the argument that ontological classification or categorisation is overrated in terms of its value in the world. He claims that experts studying a particular domain have structured traditional classification using a hierarchical taxonomy. Therefore, these systems do not satisfy the user's way of thinking and organising the world. Meanwhile, Gruber (2007) criticises Shirky's approach in that he fails to point out that folksonomies have limitations to represent, share, exchange, and reuse tags, and confuses *ontology-as-specified-conceptualisation* with a very narrow form of specification. Hendler (2007) also argues that Shirky missed the point on how ontologies could be built and what the Semantic Web is about. Spivack (2005) argues that folksonomies are just special, highly simplistic cases of ontologies with little semantics.

## The Semantic Web

The Semantic Web created by Tim Berners-Lee aims to create a universal medium for the exchange and the integration of documents and data shared and processed by software agents as well as people. The original vision of the Semantic Web appeared in 1998 (Berners-Lee, 1998) and a *Scientific American* article (Berners-Lee *et al*., 2001) elaborated the definition and the use of cases. In this article they introduced the evolution of the Web with many scenarios, such as scheduling appointments, finding documents, and locating services.

From a tagging point of view, the Semantic Web technologies can be regarded as a complement to folksonomies. While a user may interpret a tag's semantics through using or reading it, computers cannot automatically understand the meaning, since it is not defined in

a machine-readable way (Passant & Laublet, 2008). At the basic level, these technologies allow tagging data to be represented by an appropriate semantics. In particular, ontologies offer a common conceptualisation of a domain via expressing agreements on meaning. The Semantic Web also provides a common framework that allows tagging data to be shared and reused across applications or community boundaries (TagCommons, 2007a). In short, the following three areas are the main benefits of the combination of folksonomies and semantic technologies:

- *Knowledge Representation Sophistication*: Ontologies can robustly represent tagging entities and relationships among them that shape tagging activities. It could make the knowledge structure of tagging data explicit and facilitate the Linked Data (Berners-Lee, 2006) of tagging data on the Web.
- *Facilitation of Knowledge Exchange*: Ontologies enable knowledge exchange among different users and applications by providing reusable constructs. Thus, ontology for tagging can be shared and used for separate tagging activities on different platforms.
- *Machine-processable*: Ontologies and Semantic Web technologies in general (knowledge representation, processing, and reasoning) expose human knowledge to machines in order to perform automatic data linking and integration of tagging data.

## Tag ontology

There are many attempts to overcome a folksonomy's limitations by using Semantic Web technologies. Those studies detail the benefits of using Semantic Web technologies on folksonomies. Gruber (2007) and Spivack (2005) emphasise the importance of folksonomies and ontologies working together. Gruber proposes the 'Tag Ontology', which is to identify and formalise a conceptualisation of the activity of tagging, and building technology that commits to the ontology at a semantic level. The core concept in this model is *Tagging* that is the act of associating tags with an object or item (Gruber, 2007). The Tagging is comprised of the core concepts as follows:

$$\textbf{Tagging:}\ (Tag,\ Object,\ Tagger,\ Source,\ Polarity) \quad (1)$$

where *Object* refers to a thing to be tagged, identifiable by a URL or a similar naming service, *Source* refers to a scope of namespaces or universe of quantification for the object, *Polarity* is a vote for or against the assertion of the tagging. His model can be considered as a first step towards a general applicable representation model for tagging, even if his model itself is not ontology. It clearly reveals a generic conceptualisation of tagging.

Newman's model (Newman *et al*., 2005) describes relationships between an agent, an arbitrary resource, and one or more tags. In this model, there are three core concepts such as *Tagger*, *Tagging*, and *Tag* to represent a tagging activity. MOAT (Meaning of a Tag) is intended for semantic annotation of content by providing a meaning for free-text tagging (Passant & Laublet, 2008). In addition to extensions to the *Tag*, *Tagging*, and *Tagger* concepts from Newman's ontology, MOAT provides the Meaning class to represent custom, user-provided 'meanings' for tags. This class allows the meaning of tags to be made unambiguous. The dedicated approaches are focused on tagging activities or events that people used to tags in resources using terms. Therefore, the core concept is *Tagging*. The concept of tagging has a relationship, as a concept, with *Tagger* and *Object* to describe people who participate in a tagging event and objects to where a tag is assigned.

To summarise, there is an agreement on the issue as to what are the most elementary building blocks of a model for tagging. The building blocks consist of the taggers, the tags themselves, and the resources being tagged (Kim *et al*., 2008b). However, a folksonomy model needs to reflect comprehensive characteristics of tagging activity (Kim *et al*., 2008b). In the following section we discuss a common conceptualisation of tagging activity, including existing models and a proposed model that is extended.

## Conceptualising tagging and folksonomies

A tagging model needs to distinguish between entities in a tagging activity, and to address the relationships that exist between them. After reviewing existing tagging models, we discuss whether the proposed models are suitable to represent collaborative tagging activities. We then propose our extended model, which caters for the collaborative aspect of folksonomies.

### A model for tagging activities

Many researchers (Cattuto *et al*., 2007; Halpin *et al*., 2007; Mika, 2007) suggested a *tripartite* model of tagging activities, which consists of tagging entities (i.e., users, tags, and resources):

$$\textbf{Tagging:}\ (U,\ T,\ R) \quad (2)$$

where $U$ is the set of users who participate in a tagging activity, $T$ is the set of available tags, and $R$ is the set of resources being tagged. Gruber (2005) suggested an extension to model (2):

$$\textbf{Tagging:}\ (tagger,\ tag,\ object,\ source,\ +or-) \quad (3)$$

where *object*, *tag*, and *tagger* correspond to $R$, $T$, and $U$ in the tripartite model. The *source* refers to the tag space where the tagger applies the set of tags, whereas the positive/negative parameter is an attempt to represent the collaborative filtering of 'bad' tags from spammers. This tagging model has successfully been used for representing the tagging process at a semantic level. In fact, most tag ontologies have a Tagging class, based on Gruber's model, as a core concept.

### A model for collaborative tagging activities

Existing models consider tagging as an activity where an individual user assigns a set of tags to a resource. While they provide effective ways to describe the tagging process, they do not really support collective tagging activities. We therefore want to provide a *Folksonomy Model* to represent this knowledge, where the folksonomy is considered as a collection of instances of the tagging model. Before doing so, we need to clarify the differences between simple (individual) and folksonomy-based tagging practices. Folksonomies are not created independently by individuals in isolation, but collectively by people who participate in the collaborative tagging activity. Thus, the folksonomy model has to cover all the collaborative aspects and relationships in addition to the objects associated with tagging activities. A straightforward model for a folksonomy could be defined as follows:

$$\textbf{Folksonomy:}\ (\textit{user group}, \textit{tag set}, \textit{source}, \textit{occurrence}) \tag{4}$$

where the tag set is the set of all tags being employed, the user group is a set of users who participate in the tagging activity, and the source is the location where the folksonomy is utilised (e.g., social websites, online communities). The fourth parameter, *occurrence*, plays an important role to identify the tags' popularity. Comparing this model to the tagging model (2), we can identify the following similarities: the resources (objects) are not part of the Folksonomy model *per se*. The folksonomy is rather applied to the collective tagging process of the resources. The tag and tagger parameters in (2) have been replaced with a collective representation of these entities – tag set and user group. The source is still unique since a folksonomy is a multi-user approach to tagging on a single platform. In our opinion, filtering should not be represented at this level. Alternatively, given we represent multiple tags in this model, the frequencies of individual tags become important. Thus, we include the occurrence as our fourth parameter.

Contrary to the concept of Tagging, a folksonomy is a collection of individual tagging instances. It can be considered as the practice of acquiring knowledge from collaborative tagging processes. In practice, this means that the Folksonomy model should include a representation of the collective tagging processes performed by the group of users. We reflect this in (5) by extending (4) to make the individual tagging activities (to which single users contribute) explicit:

$$\textbf{Folksonomy:}\ (\textit{user group}, \textit{tag set}, \textit{source}, \textit{occurrence}, \textit{Tagging}) \tag{5}$$

where the last parameter reflects the collective tagging processes performed by the users of the folksonomy. Thus, our Folksonomy model (5) now incorporates a representation for the collective tagging processes that are individually defined by the Tagging model (2).

### SCOT: folksonomy on the Semantic Web

SCOT is an acronym for 'Social Semantic Cloud of Tags'. The name was chosen to emphasise the goal of providing a consistent framework for expressing tagging behaviours in machine-understandable way (Kim *et al*., 2008d). This ontology aims to describe folksonomic characteristics and to offer semantic links of tagging data across heterogeneous sources. SCOT offers a collection of basic terms to describe tagging entities and their relationships, and enables social interoperability for sharing and reusing semantic tag metadata across different sources. The model also describes the properties of the tags, including their occurrence frequencies, what other tags they are used in conjunction with, and what tags they are related to through user and community usage (Kim *et al*., 2008c).

In general, a user or a community may have a number of tagging events with arbitrary relationships in between. As the user continues his or her tagging activity, the relationships between tagging entities (e.g., occurrence of tags, co-occurring tags, etc.) should then be updated. Figure 1 shows core classes and properties to describe folksonomic data. In the following, the classes are introduced.

SCOT incorporates and reuses existing vocabularies as much as possible in order to avoid redundancies and to enable the use of richer metadata descriptions for specific domains. Classes and properties from other ontologies can be used together with SCOT. During the SCOT ontology design process, some external classes and properties were identified that are suitable for reuse. Such concepts are not included inside SCOT but are used directly together with terms from SCOT to describe tagging activities (see Table 1).

### Core concepts

The TagCloud class is a specific type container grouping metadata relevant to tagging practices in online communities. In general, the term 'tag cloud' is used to realise a folksonomy in the real world and is a typical method for a visual depiction of tags. In SCOT, this term as a concept aims to describe overall social aspects of tagging activities, rather than a visualisation method. Thus, this class, as a subclass of `sioc:Container`, connects the basic components such as users, tags, and resources. For instance, the property `scot:hasUsergroup` describes users who participate in a particular tagging activity. This property allows for containing multiple users, since the range of this is `sioc:Usergroup`. Tagger (user) information is represented using SIOC.

The scot:composedOf property describes a part of a TagCloud, especially if a TagCloud consists of more than two TagClouds: the property identifies each one with the specific URL. This class offers metadata information

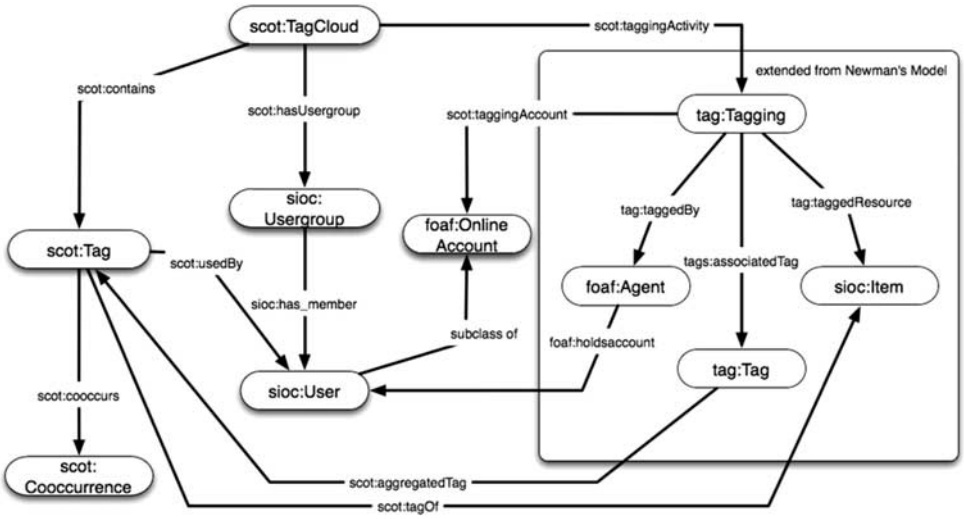


**Figure 1** Core classes and properties of SCOT. The tagging instances are represented by tag:Tagging class of Newman's model.

**Table 1 External classes and properties from existing vocabularies**

| *Prefix* | *Specification* | *XML Namespace* |
|---|---|---|
| tag | Newman's Tag Ontology | http://www.holygoat.co.uk/owl/redwood/0.1/tags/ |
| dcterms | Dublin Core Metadata Terms | http://purl.org/dc/terms/ |
| foaf | Friend of a Friend (FOAF) | http://xmlns.com/foaf/0.1/ |
| sioc | Semantically Interlinked Online Community | http://rdfs.org/sioc/ns# |
| skos | Simple Knowledge Organisation System | http://www.w3.org/2004/02/skos/core# |
| xsd | XML Schema (Datatypes) | http://www.w3.org/2001/XMLSchema# |

such as when the TagCloud generated (`dcterms:created`), where the tagging occurred (`scot:tagSpace`), and how many tags (`scot:totalTags`) and posts (`scot:totalItems`) in the TagCloud have (see Figure 2).

The Tag class, as a subclass of tag:Tag from Newman's ontology, describes a semantics of tags, which are aggregated from individual tagging activities. The Tag class is linked to `scot:TagCloud` via the `scot:contains` property and the property `scot:aggregatedTag` connects scot:Tag to a set of tag:Tag in tagging instances. There are some well-known limitations in using free-tagging classification such as 'tags' variations', 'tags' ambiguity', and 'flat organisation of tags' (Mathes, 2004; Marlow *et al.*, 2006). The limitations are critical barriers for more precise categorisation and better navigation. Some properties aim to design for solving these problems. `scot:spellingVariant` is to represent a variation in the way in which a word is spelt. `scot:delimited` is to describe a multiple-word tag name where each word is separated by a certain character. `scot:synonym` is to describe a term, which means the same as another word. Table 2 shows details of some properties for the Tag class.

These properties can reduce tag ambiguity from different conventions and even recommend more common patterns of tag name. Furthermore, in order to represent both formats of tag frequencies, SCOT introduces the two properties: `scot:ownAFrequency` and `scot:ownRFrequency`. The former is intended to describe the absolute format of popularity for a specific tag and the purpose of the latter is to represent the relative significance of the tag proportional to total tags. A single tag can have both frequency formats. Figure 3 illustrates the Tag class including relevant properties.

The Co-occurrence class defines co-occurring tags, which refer to two or more tags assigned to one resource. Generally, users assign some tags to a resource, and the tags that co-occur more frequently tend to be related. When a certain tag may often appear together with other tags, the meaning of the tag becomes more specific. For example, people can assume the tag '*apple*' is supposed to mean the company rather than a fruit because of its context – if the tag is used with the tag '*iPhone*'. A tag co-occurrence is described by using `scot:cooccursIn` or `scot:cooccurredWith` properties. The `scot:cooccurrenceAFrequency` and `scot:cooccurrenceRFrequency` properties describe absolute and relative values for co-occurring tags (see Figure 4).

These classes in SCOT play a role in the representation of the social and semantic context of tagging, since they include users, tags, and resources, and additional information to clarify tags' semantics, respectively.

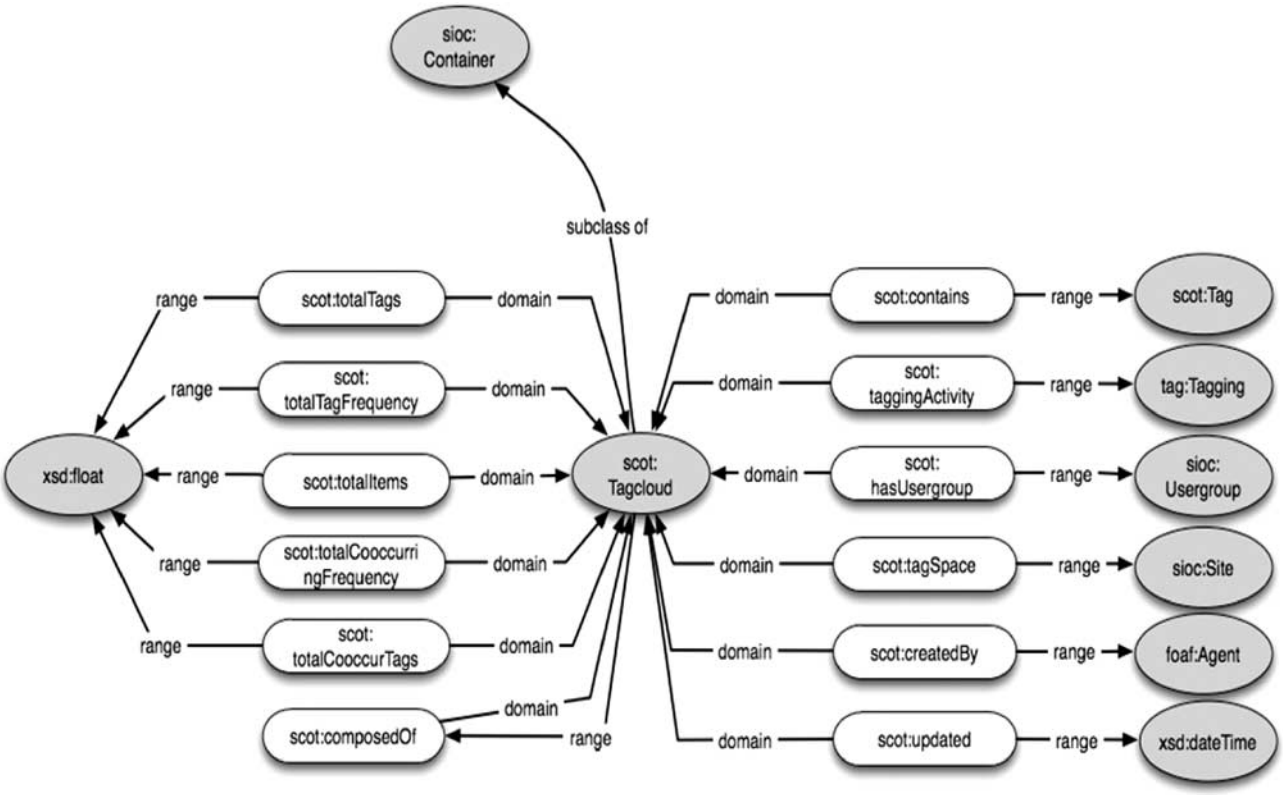


**Figure 2** TagCloud class and its properties. scot:taggingActivity is used for linking between Tagging instance and TagCloud.

**Table 2 Properties of the tag class**

| *Properties* | *Description* |
|---|---|
| scot:tagOf | Indicates that a tag is assigned to the sioc:Item |
| scot:hasTag | An item has one and more tags |
| scot:usedBy | Refers to the sioc:User who uses or creates a tag |
| scot:spellingVariant | A variation in the way in which a word is spelt |
| scot:acronym | An abbreviation formed by the first letters of the compound word |
| scot:synonym | A word that means the same as another word |
| scot:plural | A tag name refers to more than one thing |
| scot:singular | A tag name refers to one thing |
| scot:delimited | A multiple-word tag name where each word is separated by a certain character |
| scot:hypenated | A multiple-word tag name where each word is separated by a hyphen |
| scot:underscored | A multiple-word tag name where each word is separated by an underscore |
| scot:slashed | A multiple-word tag name where each word is separated by a slash |
| scot:spaced | A multiple-word tag name where each piece of word is separated by a space |
| scot:ownAFrequency | An absolute frequency of a particular tag in a TagCloud |
| scot:ownRFrequency | A percentage frequency of a tag for a particular TagCloud relative to the total of all Tag frequencies in that TagCloud |
| scot:lastUsed | The last date on which a particular tag was used |
| scot:aggregatedTag | Similar but distinct tags from different tag:Tagging are aggregated to a common tag representation in a TagCloud |

## Describing tagging activity

The Tagging class represents tags themselves, the resources that are being tagged, and the users that create these tags `tag:taggedBy`. SCOT uses Newman's ontology to describe tagging instances. As shown in Figure 5, the `scot:TagCloud` class connects `tag:Tagging` instances via the property `scot:taggingActivity`. This property describes a relationship between an instance of a `scot:TagCloud` and an instance of a `tag:Tagging`. In this way, all tagging events are collectively linked to an instance

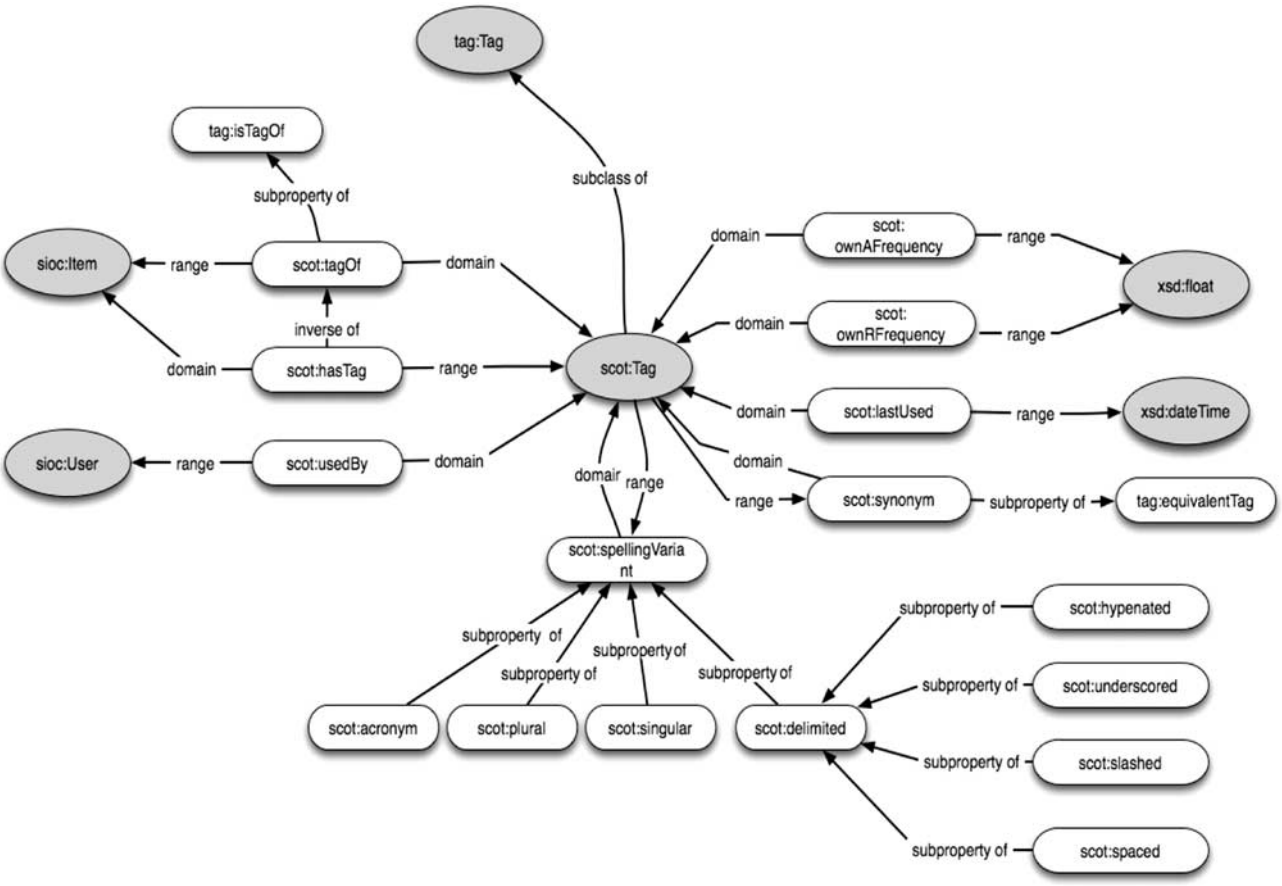


**Figure 3** Tag class and its properties.

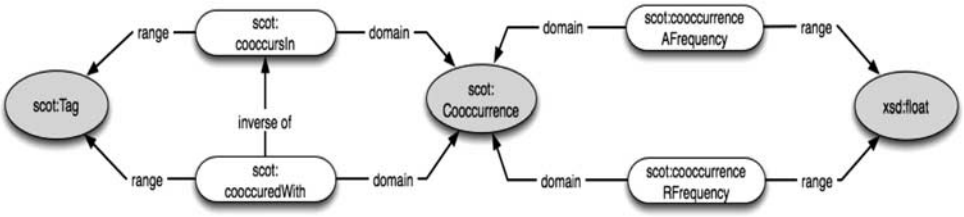


**Figure 4** The Co-occurrence class and its properties.

of the TagCloud class. SCOT also introduces the `scot:taggingAccount` describing the relationships between a tagging activity and the account used when performing the tagging. Individual tags in `tag:Tagging` are mapped to a resource with `tag:Tag` instance and then these tags are represented by a collection of tags underlying a `scot:TagCloud` via `scot:aggregatedTag`. Using `scot:tagOf` property, each tag is linked to `sioc:Item` in a tagging instance.

Figure 5 shows the merged tag space from both Bob and Alice's personal tag clouds. This approach shows how a user-centric folksonomy to represent the interests of small groups or communities can be created in SCOT. It is also possible to adopt this example across sites or resources. Although both users have tagging data on different resources, the entities of tagging activities are explicitly linked to each other and the structure of the tagging data is consistent to share and reuse. Therefore, this approach can be adopted to create their customised folksonomy. In addition, SCOT can be utilised by SPARQL, the query language for Semantic Web data.

In Figure 5, Alice and Bob use the tag 'web' to identify the resource 'Video#id'. Tagging class represents individual activities of both users, and the TagCloud class contains all users, tags, and resources in the given example. Bob and Alice's linked tag space. Both users' tag clouds are linked using `scot:composedOf` property. Figure 6 shows a snippet represented by using RDF/XML for Figure 5.

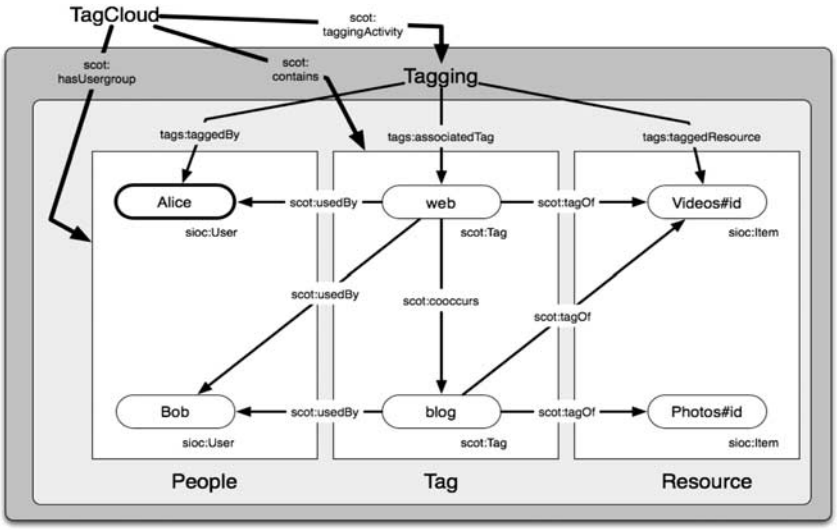


**Figure 5** The SCOT ontology and other vocabularies.

```
<scot:TagCloud rdf:about="http://www.blogweb.co.kr/sonagi/scot/tagcloud/1">
      <dc:publisher rdf:resource="http://scot-project.org"/>
    <dcterms:created>2008-11-30T17:31:49</dcterms:created>
    <scot:contains
rdf:resource="http://www.blogweb.co.kr/sonagi/scot/tag/blog"/>
    <scot:createdBy rdf:resource="http://www.blogweb.co.kr/?author=1"/>
    <scot:hasUsergroup rdf:parseType="Resource">
            <sioc:has_member rdf:resource="http://delicious.com/bob/"/>
            <sioc:has_member
rdf:resource="http://www.flickr.com/photos/alice"/>
       <rdf:type rdf:resource="http://rdfs.org/sioc/ns#Usergroup"/>
    </scot:hasUsergroup>
    <scot:tagSpace rdf:resource="http://www.blogweb.co.kr"/>
    <scot:taggingActivity
rdf:resource="http://www.blogweb.co.kr/?p=1#tagging"/>
</scot:TagCloud>
<scot:Tag rdf:about="http://www.blogweb.co.kr/sonagi/scot/tag/blog">
      <scot:lastUsed>2008-12-01T02:31:47</scot:lastUsed>
    <scot:ownAFrequency>39</scot:ownAFrequency>
    <scot:ownRFrequency>9.80</scot:ownRFrequency>
    <tag:name>blog</tag:name>
    <rdf:type rdf:resource="http://scot-project.org/scot/ns#Tag"/>
</scot:Tag>
```

**Figure 6** Example of SCOT instance metadata in RDF/XML.

## Conclusions

This paper presented a semantic model for representing the structures and the relationships between tagging entities created by collective tagging activities. Folksonomies are popular nowadays as a collaborative and collective way of categorising web resources. The problems that have been noticed in current folksonomies, however, include lack of reuse of tags between an individual's various social media applications or between communities, inconsistent use of tags, the reluctance to recreate tag sets on new systems, and the lack of an expressive format for describing tag structures and relationships.

Ontology-based Semantic Web technologies provide standards to interlink diverse folksonomic platforms and to facilitate connections between tagging data in data silos, making it available for sharing and processing by people, communities and machines. As the tags can be combined with meaningful concepts in ontologies, the basic level variation problems encountered in conventional systems will be addressed. The semantically enriched tag data can be reused or exchanged across services or users.

The SCOT ontology designed to solve these lacks of semantics of folksonomies, and developed by the machine-processable format via combining with existing RDF vocabularies. Using this model can offer new opportunities for sharing and disseminating tag metadata across different applications, communities, or websites.

## References


BERNERS-LEE T (1998) Semantic web road map. World Wide Web Design Issues [WWW document] http://www.w3.org/DesignIssues/Semantic.html (accessed 20 January 2010).

BERNERS-LEE T (2006) Linked data. World Wide Web Design Issues [WWW document] http://www.w3.org/DesignIssues/LinkedData.html (accessed 20 January 2010).

BERNERS-LEE T, HENDLER J and LASSILA O (2001) The semantic web. *Scientific American* **284(5),** 34–43 [WWW document] http://www.sciam.com/article.cfm?id=the-semantic-web (accessed 20 January 2010).

CATTUTO C, SCHMITZ C, BALDASSARRI A, SERVEDIO VDP, LORETO V, HOTHO A, GRAHL M and STUMME G (2007) Network properties of folksonomies. *AI Communications Journal, Special Issue on 'Network Analysis in Natural Sciences and Engineering'* **20(4),** 245–262.

GRUBER T (2005) TagOntology – a way to agree on the semantics of tagging data [WWW document] http://tomgruber.org/writing/tagontology-tagcamp-talk.pdf (accessed 20 January 2010).

GRUBER T (2007) Ontology of folksonomy: a mash-up of apples and oranges. *International Journal on Semantic Web & Information Systems* **3(2),** 1–11.

GRUBER T (2008) Collective knowledge systems: where the social web meets the semantic web. *Journal of Web Semantics* **6(1),** 4–13.

HALPIN H, ROBU V and SHEPHERD H (2007) The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, pp 211–220, Banff, Alberta, Canada, 08–12 May 2007. WWW'07.ACM, New York, http://portal.acm.org/citation.cfm?id=1242572.1242602.

HENDLER J (2007) Shirkyng my responsibility. [WWW document] http://www.mindswap.org/blog/2007/11/21/shirkyng-my-responsibility/ (accessed 20 January 2010).

HENDLER J, SHADBOLT N, HALL W, BERNERS-LEE T and WEITZNER D (2008) Web science: an interdisciplinary approach to understanding the web. *Communications of the ACM* **51(7),** 60–69.

KIM HL, BRESLIN JG, YANG SK and KIM HG (2008a) int.ere.st: building a tag sharing service with the SCOT ontology. In *Proceedings of AAAI 2008 Spring Symposium on Social Information Processing* (LERMAN K, GUTELIUS D, HUBERMAN B and MERUGU S, Eds), Technical Report SS-08-06. 2008.

KIM HL, PASSANT A, BRESLIN JG, SCERRI S and DECKER S (2008b) Review and alignment of tag ontologies for semantically-linked data in collaborative tagging spaces. In *Proceedings of IEEE International Conference on Semantic Computing* pp, 315–322, IEEE, USA, http://portal.acm.org/citation.cfm?id=1446386.

KIM HL, SCERRI S, BRESLIN JG, DECKER S and KIM HG (2008c) The state of the art in tag ontologies: a semantic model for tagging and folksonomies. In *Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications*, pp. 128–137, Berlin, Germany, 22–28 September 2008. Dublin Core Metadata Initiative, http://portal.acm.org/citation.cfm?id=1503418.1503431.

KIM HL, BRESLIN JG, SCERRI S, DECKER S, KIM HG and YANG SK (2008d) SCOT ontology specification [WWW document] http://scot-project.org/scot/index.html (accessed 20 January 2010).

KROSKI E (2005) The hive mind: folksonomies and user-based tagging [WWW document] http://infotangle.blogsome.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/ (accessed 20 January 2010).

MARLOW C, NAAMAN M, BOYD D and DAVIS M (2006) HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*, pp 31–40, Odense, Denmark, 22–25 August 2006. HYPERTEXT '06, ACM, New York, http://portal.acm.org/citation.cfm?id=1149941.11449949.

MATHES A (2004) Folksonomies – cooperative classification and communication through shared metadata [WWW document] http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html (accessed 20 January 2010).

MERHOLZ P (2004) Metadata for the masses [WWW document] http://adaptivepath.com/publications/essays/archives/000361.php (accessed 20 January 2010).

MIKA P (2007) Ontologies are us: a unified model of social networks and semantics. *Journal of Web Semantics* **5(1),** 5–15.

NADLER L (2006) Gartner highlights seven core benefits of Web 2.0 for traditional industries, http://www.gartner.com/it/page.jsp?id=499154 (accessed 20 January 2010).

NEWMAN R, AYERS D and RUSSELL S (2005) Tag ontology [WWW document] http://www.holygoat.co.uk/owl/redwood/0.1/tags/ (accessed 20 January 2010).

PASSANT A and LAUBLET P (2008) Meaning of a tag: a collaborative approach to bridge the gap between tagging and linked data. In *Proceedings of the Linked Data on the Web Workshop (LDOW2008)* (BIZER C, HEATH T, IDEHEN K and BEMERS-LEE T, Eds), 22 April, Beijing, China, CEUR Workshop Proceedings, ISSN 1613-0073, online CEUR-WS.org/Vol-369/.

QUINTARELLI E (2005) Folksonomies: power to the people [WWW document] http://www.iskoi.org/doc/folksonomies.htm (accessed 20 January 2010).

RAINIE L (2007) 28% of online Americans have used the internet to tag content [WWW document] http://www.pewinternet.org/pdfs/PIP_Tagging.pdf (accessed 20 January 2010).

SHIRKY C (2005) Ontology is overrated: categories, links and tags [WWW document] http://shirky.com/writings/ontology_overrated.html (accessed 20 January 2010).

SPIVACK N (2005) Folktologies-beyond the Folksonomy vs. ontology distinction [WWW document] http://novaspivack.typepad.com/nova_spivacks_weblog/2005/01/whats_after_fol.html (accessed 20 January 2010).

TAGCOMMONS (2007a) Functional requirements for sharing tag data [WWW document] http://tagcommons.org/2007/02/28/functional-requirements-for-sharing-tag-data/ (accessed 20 January 2010).

TAGCOMMONS (2007b) Ontologies Vs. formats Vs. schema Vs. APIs [WWW document] http://tagcommons.org/2007/03/02/ontologies-vs-formats-vs-schema-vs-apis/ (accessed 20 January 2010).

VANDER WAL (2005) Folksonomy definition and wikipedia [WWW document] http://www.vanderwal.net/random/entrysel.php?blog=1750 (accessed 20 January 2010).

WEINBERGER D (2005) Taxonomies and tags: from trees to piles of leaves. *Esther Dysons Monthly Report* **23(2),** 1–36.


## About the authors

**Hak-Lae Kim** is currently a senior engineer of Media Solution Center at Samsung Electronics Co., LTD. He is also a director of Web Science Workgroup in Korea. He was a researcher at the Digital Enterprise Research Institute at NUI Galway, and his research efforts focused on creating social spaces based on tagging activities. He is founder of the SCOTs ontology and project, aims to provide semantic relations for tagging data, and has implemented the int.ere.st web site for sharing tagging data. Hak-Lae Kim completed his B.S. in Business Administration, and M.S. and Ph.D in Management Information System from Dankook University, and Ph.D in Computer Science from National University of Ireland, Galway.

**John G. Breslin** is currently a lecturer in *Electrical and Electronic Engineering* at NUI Galway. He is also an associate researcher and leader of the *Social Software Unit* at the *Digital Enterprise Research Institute* at NUI Galway, researching semantically enabled social networks and community portals. He is the founder of the *SIOC project* (*Wikipedia*

*article*). He was previously a postdoctoral researcher at DERI from 2004 to 2008, a lecturer at the University from 2000 to 2004, and was a research officer and postgraduate student at the *Power Electronics Research Centre* in NUI Galway before that. He also worked as a visiting scholar at *Virginia Tech* in 1996. He has received a number of awards for website design, including a Golden Spider for the Irish community website *boards.ie* (*Wikipedia article*), which he co-founded in 2000. The Irish Internet Association presented him with Net Visionary awards in 2005 and 2006.

**Hong-Gee Kim** is an associate professor of medical informatics and the director of Biomedical Knowledge Engineering Lab at Seoul National University. He also holds an adjunct professorship in the college of engineering and informatics at the National University of Ireland, Galway. He was previously a professor of MIS in a business school. His primary research efforts are in building real world applications by using ontology and semantic technologies in the e-Health domain. He is currently leading various research projects concerning a semantic terminology server for electronic health record systems, a semantic search engine on the cloud computing environment, and a social semantic digital library system. He published approximately 100 academic papers. Hong-Gee Kim completed his M.S. and Ph.D. in Artificial Intelligence and Philosophy from the University of Georgia and his B.S. in Psychology from Korea University.

**Jae-Hwa Choi** is a professor of Information Systems area in the Department of Management at the Dankook University in the Republic of Korea. His recent research interest includes collaborative innovation and collective intelligence in knowledge management and business intelligence and decision support systems.