This paper was published in:

The Proceedings of the iConference 2011, Seattle, WA, February 8-11, 2011

ICPS: ACM International Conference Proceeding Series

An Emergent Culture Model for Discerning Tag Semantics in Folksonomies

David J. Saab

College of Information Sciences and Technology
The Pennsylvania State University
University Park, PA 16802

dsaab@ist.psu.edu

ABSTRACT

Social bookmarking sites as Flickr, del.icio.us, and CiteULike have incorporated the use of tags as way for users to retrieve photos, URLs, and citations in a way that is personally meaningful and which doesn't require learning taxonomies constructed by professionals. These tag sets, or folksonomies, have the potential to enhance interoperability among our information systems, especially those that use computational ontologies. Formal computational ontologies form the foundation for semantic interoperability, but seem to be insufficient in facilitating it because the ontologies developed for different information systems do not have an inherent mechanism for negotiating meaning or recognizing the natural evolution of a lexicon. Coupling folksonomies with formal ontologies holds potential for more productive semantic interoperability among systems. In order to reach that potential, we need to understand more clearly the process of discerning semantics in tag sets as entry points into the complex conceptual networks that generate meaning within cognition. This paper will explore that semantics involved in "emergent semantics" in tag sets and introduce an emergent culture model that will help clarify how folksonomies can be utilized in this endeavor.

Categories and Subject Descriptors

H.1.1 [**Information Systems**]: Models and Principles—Systems and Information Theory (Value of information)

General Terms

Theory, Design

Keywords

Folksonomies, tags, tagging, cultural, schemas, ontology, semantics, hermeneutic

1 SEMANTIC INTEROPERABILITY

Formal ontologies, upon which computational ontologies are modeled, are types of categorization systems that are used by philosophers and computer scientists to describe what exists in reality. Ontology is a *philosophia prima* concerned with the theory of being, i.e., what exists. In his Metaphysics, Aristotle describes Ontology as regarding "all the species of being qua being and the attributes which belong to it qua being" [2]. A "true" ontology would be one—and there would be only one—in which all things of existence and their relationships with one another were described in a single coherent and comprehensive treatise [36]. Formal ontologies are a complex form of metadata that specify the underlying concepts and their relationships that comprise the information *of* and *for* an information system [15]. The most common understanding of ontology in computer and information sciences is Gruber's "specification of a conceptualization" [6, 19]. However, formal ontologies are problematic in that they simultaneously crystallize and decontextualize information, which in order to be meaningful must be adaptive in context. In trying to construct a correct taxonomical system, formal ontologies are focused on syntactic precision rather than meaningful exchange of information and semantic interoperability.

Folksonomies is a term coined by Vander Wal [43] to refer to the "result of personal free tagging of information and objects for one's own retrieval." Tagging happens in a social environment and is done by individuals consuming information. Tags are generated by individuals for their personal use, to be able to retrieve information and/or objects quickly and in a way that conforms to their understanding of the entity. Social bookmarking sites as Flickr, del.icio.us, and CiteULike have incorporated the use of tags as way for users to retrieve photos, URLs, and citations in a way that is personally meaningful and which doesn't require learning taxonomies constructed by professionals. Users employ their own vocabulary, which has meaning specific to them. It is these meaningful associations expressed as tags that enable faster and more direct recall of the object because they act as representations for the way we think [21]. However, in order for these meaningful associations to be useful in a sociocultural context, they must be aligned conceptually so that their semantics remain consistent across tag sets.

Researchers have made various efforts, with limited success, at coupling ontologies and folksonomies to derive some sort of semantic interoperability [4, 18, 20]. The difficulty with this coupling reveals a multifaceted tension between 1) structured ontologies and unstructured folksonomies, 2) lexicality, syntax and semantics, 3) social and cultural dimensions of tags, and 4) emergence and reification. In this paper I will explore these

issues as they relate to the potential for an emergent semantics derived from folksonomies through the use of an Emergent Culture Model (ECM) I developed. The explanatory power of the ECM is limited to the enactive dimensions of tags and tag creation. It is best utilized as a means of structuring the relationships between the smaller network of tags (often cast as semantic networks) and the larger cognitive-cultural landscape in which semantics are actually discerned [31]. The intent behind the ECM is to provide a relatively sophisticated mechanism for a more nuanced understanding of tags and how they work, which hopefully leads to higher quality research in relation to tags.

2   THE CULTURAL NATURE OF TAGS

Tags at their most basic are simple lexical units (i.e., words). They can be more complex as when combinations of words are strung together in various ways (e.g., informationvisualization, infoviz, user-generated-content or semantic_web)—assuming that tags are determined by space separation and not comma separation. Tags that are aggregated into tag sets or tag clouds are simply a collection of lexical units. Such aggregations might consist of the tags produced by a single individual or a group of individuals. They might consist of tags aggregated by particular resources (i.e., what is signified by the tag) such as a URL in del.icio.us, an image in Flickr, or a citation in CiteULike. It is through aggregating tags that researchers imagine semantics might emerge [3, 5, 8, 32, 33, 35, 38].

Researchers sometimes mistake the collective nature of tagging and creating folksonomies with a collaborative activity, which leads to misunderstanding of the social and cultural nature of tags [32]. The social nature of tags is reflected in language and the shared vocabulary that we use to convey semantics. It is this social activity of sharing vocabulary that gives rise to power law distributions of tag sets. A word isolated from the entity it was intended to describe and from the person who created it can mean or refer to many things, and many people may interpret the same tag differently based on their personal histories. Tags are ontic signs that serve as indicators to the rich ontological conceptualizations we hold in cognition. In order for the semantics of tag sets to emerge, it is important to understand the perspective from which the tags are offered. Because each individual has a different experiential history, we would expect that their ontological conceptualizations to be unique. However, individuals are also members of cultures, and as cultural beings they share many common experiences and articulate them using language.

Culture, as described here, is an emergent phenomenon arising through the interplay of patterns within cognition with patterns extant in the world we inhabit. Schemas, as complex cognitive associations, are intrapersonal structures. The objects or events that are manifest outside individual cognition, the entities in the external world, are extrapersonal structures. Culture consists of the interplay between the intrapersonal cognitive structures and extrapersonal structures such as systems of signs, infrastructure, environment, social interaction, and so on [40]. It is through this interplay that we can see that some of the intrapersonal cognitive structures called schemas are shared with others, making them cultural schemas [40].

This characterization of culture is somewhat at odds with our linguistic conventions and might seem strange. In common speech, we refer to culture as if it were an entity separate and

apart from human cognition. I can refer to Western culture, or Brazilian culture, or indigenous cultures, and these would have meaning for us. When we talk of culture as a thing rather than a phenomenological process in terms of schemas, however, we often mistake culture for belief. As culture is defined here, it is an emergent phenomenon, not some thing that exists independently of human experience. Culture is not only cognitive or simply a set of beliefs or norms, it is the *interplay*. When schemas are shared, they become *cultural*, in the sense of distributed cognition [24]. Because culture is defined this way, when we talk about atheist culture or socialist culture or photographer culture as well as Western, Brazilian and indigenous cultures, we are actually talking about cultural schemas enacted phenomenologically and at various levels of abstraction [34]. The logical extension of this way of defining culture is that individuals become "junction points for an infinite series of overlapping cultures" [40].

In cognitive science, connectionist theory posits the human conceptual system as a network composed of a large number of units joined together in a pattern of connections [30]. Cognitive scientists, cognitive anthropologists, cognitive neuroscientists, and educational psychologists refer to these patterns of connections as schemas [1, 7, 9-11, 25, 39, 40], but they have also been referred to variously in the literature as frames, scenes, scenarios, scripts, models, and theories [9]. For my purposes here, schemas will refer to these patterns generally, and frames (in section 3) will refer to a specific subset of schemas tied to a particular cultural identity that constrains the semantics of tags. Schemas are strongly connected networks of cognitive elements, having a bias in activation through repeated exposure to the same or similar stimulus, but they are not rigid and inflexible. They are adaptable, sometimes resulting in the strengthening of existing schemas, sometimes in their weakening in the face of new experience. Schemas are "flexible configurations, mirroring the regularities of experience, providing automatic completion of missing components, automatically generalizing from the past, but also continually in modification, continually adapting to reflect the current state of affairs" [9]. In other words, schemas are associative in nature and based in personal historical experience.

Describing them as 'flexible, mirrored configurations' implies that schemas are structural entities within cognition that are comprised of several elements. Schemas are not the individual elements rather strongly connected clusters of elements of experience within cognition. Elements of experience are clustered in cognition because they are clustered in our lived experiences. Clustering cognitive elements makes them more efficient by reducing the cognitive load associated with processing experience. Because of the strong connections of elements and their clustering, we are able to fill in information that may be missing from our experience. Schooling, for example, has a number of associated elements—teachers, peers, desks, books, reading, writing, lecturing, socialization, athletics, and so on. When someone mentions 'school' or 'schooling', there are a variety of schemas evoked in cognition based upon our lived experience with the same. The schemas evoked in a teacher would be different than the schemas evoked in a student. Simply mentioning the name of something (or reading it as a tag) is often enough to activate schemas associated with it. Schemas help to fill in the ambiguous or missing information because the associated neurons are more likely to be activated by the initial stimuli.

Schemas are powerful processors of experience, help with pattern completion, and promote cognitive efficiency. They serve to both inform and constrain our understanding of experience. People recall schematically embedded information more quickly and more accurately [12]. In fact, schemas hold such sway in our cognition that people may falsely recall schematically embedded events that did not occur. They are more likely to recognize information embedded in existing schemas because of repeated activation of the schemas. This repeated activation evokes expectations within cognition, and the easy recognition of contradictory or challenging information that do not conform to those expectations formed as part of the existing schemas. Information that is orthogonal to existing schematic structures, that doesn't acquire salience through the repeated activation of schemas and the creation of associated expectations, is much less likely to be noticed or recalled. Because of their functionality in pattern completion, schemas function, in some sense, as flexible filters of experience, enabling us to attend to its salient features while filtering out the non-salient.

Tags are not references, per se, but rather indicators for the cognitive schemas that are activated upon encountering the tag. Heidegger might say that when we encounter a tag, as when we encounter a sign, our activated schemas make salient parts of the environment in which it is embedded, and the encounter orients us in a particular way, making us ready to engage 'what is coming.' Tags indicate where one's concern dwells, and characterizes what sort of involvement one has with something [22]. Where our concern dwells as we create a tag may not simply be as a neutral descriptor of a particular entity or phenomenon, but may also reflect meta concerns, such as critiques or attitudes, as one reviewer of this paper pointed out when he uses the tag, "dude!" Such a tag can indicate a positive or negative attitude, and its semantics cannot be discerned unless and until one determines which cultural schemas and related identities were employed in its creation. Tags form entry points into our complex of cognitive and cultural schemas that shape our ontological commitments to the world in which we are immersed.

In terms of creating tags, when we use them for personal recall, we are identifying the salient qualities and dimensions of our experience with the phenomenon or entity being tagged. From the ontological, we create the ontic sign—the tag. They are meaningful to us because they are created based on how we understand the phenomenon, which is in turn based on our personal historical context. Tags become an indicator of that salient experience. They allow us to reactivate our ontological understanding (i.e., activate our schemas) in later encounters with the tags that we create. Tags, we should remember, are not always simply created for personal recall, but sometimes as commentary on the subject about which the resource addresses. When tags are created in this vain, tags become a second-order indicator such that they deal not with the resource as the entity or phenomenon, but with that entity or phenomenon to which the resource refers and our relationship to it. If, for example, I tag something "reductionist," that could be because the resource describes something that is approached with a "reductionist" theoretical perspective. However, "reductionist" also indicates for me a lot of other senses, mostly critical and tending toward the negative, that are clustered by virtue of the schemas that I have developed in relation to the philosophy of science. Whether I create tags for personal recall or to create commentary, the tag serves as an indicator—directly or indirectly, first-order or second--order—of a cultural schema and the salient entry point into that schema's associations.

Figure 1 illustrates the emergent model of culture. It is in the cultural nature of tags where semantics can be discerned. Semantics, like culture, is an emergent phenomenon, and would emerge as part of culture in this model. Semantics requires a conceptual network, not just a lexical network of tags configured in a particular way. As ontic signs, tags are extrapersonal structures, though they have simple lexical rather than complex structural or content dimensions. It is through the interplay of the tags and the intrapersonal schemas that semantics emerge. Tags will evoke schemas as the individual interacts with them, and it is through this interaction that meaning will emerge. While lexicality and syntax are necessary elements in the emergence of semantics of tags, they are insufficient for semantic emergence.
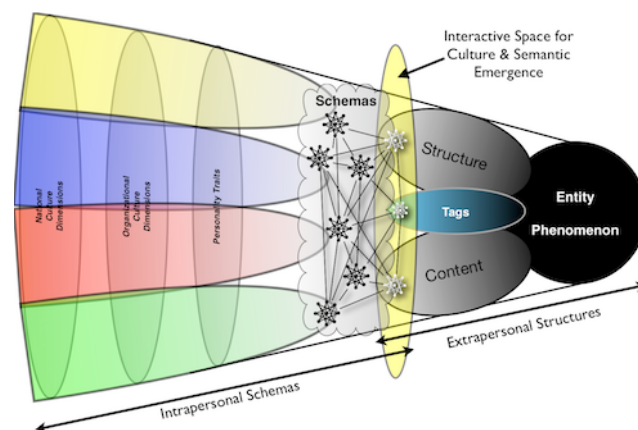

Figure 1. Emergent Culture Model

Application of the model is actually more complex than delineating boundaries around sets of individuals or sets of cultures. Identity plays a role in the emergence of culture such that individuals have multiple cultural identities. Individuals may create tags that relate to several different cultural identities. We must be able to disaggregate the tags of an individual into sets derived from that individual's cultural identities. We discuss further the relevance of cultural identity perspectives in the next section.

3 MULTIPLICITY OF IDENTITIES

Recalling the idea of individuals forming a junction point, the notion of identity and multiplicity of perspectives is important in our understanding how cultural schemas manifest. Individuals can manage multiple identities in the same or multiple contexts [41]. People living in multicultural environments often encounter situations which require them to acquire different cultural schemas and to switch between these cultural schemas depending on their immediate sociocultural context [7]. We can shift our perspective effortlessly between national, familial, peer and other identities to make sense of particular phenomena (i.e., frame it in relation to ourselves). The same context, for example, that would be considered "exciting" to "the hunter" might also be "dangerous" to "the parent." It may be the danger that excites the hunter, who as a parent wants to shield his children from such danger. The same word and the same concept of "danger" has

very different semantic implications in different contexts. Fauconnier and Turner claim that "frames structure our conceptual and social life and, in their most generic and schematic forms, create a basis for grammatical construction" [13]. Words are themselves viewed as constructions, and lexical meaning is an intricate web of connected frames. Although cognitive framing is reflected and guided by language, it is not inherently linguistic—people manipulate many more frames than for which they have words and constructions. It is the individual's salient, contextualized identity in relation to the phenomena that allows for sense making of the phenomena. When making meaning of a particular phenomenon, individuals will rely upon the cognitive and cultural schemas that are integral parts of their salient, contextualized identities. Tags may facilitate personal recall, but also they facilitate the recall of one's identity inasmuch as they facilitate recall of the tagged entity.

The ECM in Figure 1 captures the interaction between language and identity. In the model, language would be an extrapersonal structure. Identity is a particular configuration of intrapersonal schemas based on a variety of contextual factors manifesting at a variety of levels—national culture level, organizational culture level, and personality level. The difficulty in recognizing intrapersonal schemas is the same difficulty we have in recognizing cognition: we can't observe them directly. We have to interpret and infer what people are thinking via their actions and speech. And this is always done in context. Though people's intrapersonal schemas can never be shared directly, we are able to share them to some extent through language. It is through this sharing that we develop a *cultural* understanding of the patterns and connections of entities and phenomena we encounter in experience. It is these larger cultural patterns of intrapersonal schemas—cultural schemas—that we might posit as representative of identity. The mere mention of a word is often sufficient to evoke any number of schemas. As extrapersonal structures, words and language (i.e., tags) serve as social representations that help us identify relationships between images, ideas, objects, and phenomena we encounter in the world [14, 29]. They form the entry points into our complex intrapersonal schemas and rich ontological understanding of experience. What intrapersonal schemas a tag will evoke is dependent upon the cultural context in which it is being experienced [37].

A shared vocabulary is negotiated over time and evokes cultural schemas within an individual's cognition. A shared vocabulary has meaning to the cultural group because the semantics emerge through the evocation of the ontological (i.e., schemas) via the ontic (i.e., tags). The stabilization of tag patterns over time [17] is analogous to the stabilization of cognitive schemas as cultural schemas. The collective tags of a folksonomy will certainly reflect the dominant cultural schemas of a broad population, but the assumption that collective tags represent a shared (cultural) conceptualization, interferes with discerning minority cultures, whose schemas may overlap with but are not necessarily entirely consistent with those of the dominant cultural group. In the absence of perspective and cultural identity information about users, folksonomies can be considered as reflections of cultural schemas only for dominant cultural groups and only in the broadest possible sense of "cultural group" [32].

Another difficulty we face with the aggregation of tags is the fact that individuals utilize multiple cultural identities. We may share some cultural schemas but not others with particular persons. I may create different tags for the same entity based on different cultural identities and different cultural schemas, as in the example of hunter/parent above. Discerning cultural identity within and across individuals is not an easy task, as individuals may not even recognize the cultural identities they are using in creating the tags. However, if we are to discern the emergent semantics of folksonomies, we must include the perspective from which the person is creating the tag.

The ECM can be modified to represent multiple cultural identities and perspectives. In Figure 2 I illustrate the integration of the culture model using a mandala to represent eight perspectives on a phenomenon. In reality, the perspectives may overlap a great deal more or a great deal less depending on the extent of overlap of shared cultural schemas, but they are depicted as a mandala for ease of discussion, not to suggest that our cultural identities are completely independent of one another.
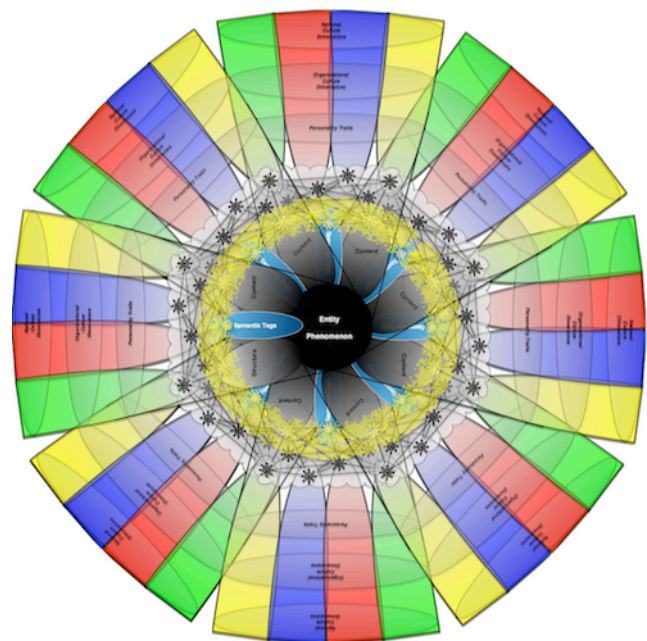

Figure 2. ECM Mandala illustrating eight perspectives on a single phenomenon.

We can contextualize the interactions between perspectives in different ways and illustrate the flexibility of the model to portray culture at individual, group, organizational and societal levels of analysis. In the context of a single culture, where the entire mandala is understood to represent a common cultural identity, each perspective would correspond to a single person. Each person would have a set of tags they created for the entity or phenomenon. Those sets of tags would be reflective of the intrapersonal schemas of a single individual and would consist of the individual's cognitive and cultural schemas. Because the tag sets are reflective of the cultural schemas of a single culture, we would predict that the individuals would use many of the same tags to describe the phenomenon. The aggregated tag sets would exhibit a power law curve, with a few particular tags with high frequency and a relatively larger number of tags with low frequency.

We could also contextualize the interactions between different perspectives such that they are intercultural, arising from different

cultural schemas. We could consider each of the eight perspectives unique to a particular cultural perspective (of individuals, of organizations, of nations, of societies, etc.). Each set of tags would be created from those different cultural perspectives and reflect a much greater diversity of conceptualizations with respect to the entity or phenomenon under examination. We might see several power law curves and see distinct clustering of sets of tags based on the different perspectives. Comparisons of the tags sets in this type of configuration might yield places where different cultures can begin a discourse that would allow for the negotiation of meaning across cultures.

The simplicity of the model belies its strength in terms of scalability. The eight perspectives illustrated in Figure 2 could be applied to the cultural identities of a single individual (e.g., male, student, New Englander, photographer, technologist, socialist, gay, atheist). The ECM could represent several individuals from any one of these cultural identities. Or, it could be used to represent multiple individuals from each of these identities. Like with Actor-Network Theory [26, 27], there is an inherent flexibility to the model in terms of scaling. A node can be representative of a single individual or a sub-network, just as the ECM mandala can represent a single individual (with multiple cultural identities) or several individuals (with the same or different cultural identities). The entity or resource that is being tagged remains a constant in any application of the model, but the perceived structure or content or the resource, as well as the tag associated with it, is dependent as much upon the cultural identity being mapped as it is to the entity to which these qualities belong. The ECM and its mandala variation allows us to properly articulate the level at which the tag analysis is being done and prevents confusion of the social and cultural dimensions of tags by ensuring that both are considered appropriately.

4   MULTIPLICITIES OF MEANING

Having disambiguated the issue of social and cultural dimensions of tags with the ECM, we need to turn our attention to the semantics of tags. When dealing with the semantic dimensions of tags, issues of polysemy and synonymy reveal themselves [17]. Tags are created at basic, superordinate, and subordinate levels and are related to an individual's interactions with them [42]. How does one disambiguate among polysemous or synonymous tags? One solution for disambiguating tags is to add a specification to OWL (Web Ontology Language) such that "`<tag> owl: DifferentFrom <tag>`", where the tag is the same lexical unit (e.g., apple) but has different meaning (e.g., fruit vs. computer company) [28]. A complementary suggestion includes the use of "`owl: SameAs`" to merge tags with the same meaning (e.g., *semweb* and *semantic web*). This approach, such as it is, looks promising, but it doesn't easily account for the evolution of the collective lexicon. It would also put a burden upon the tagger to specify the "`owl: Relationship`" in a tagging system.

Requiring taggers to specify `DifferentFrom` and `SameAs` relationships only addresses the ontic side of the equation and thre social nature of tags [32]. In order to address the ontological, our understanding of the user as part of the semiotic-like relationship must not neglect his cultural perspectives and identities when trying to discern the semantics of particular tag sets. We must consider meaning-making, which is a cultural activity, as a multifaceted process, where semantics emerge through a process of interaction, construction and communication [38]. Interaction involves tasks and activities that generate the need for new meanings based on our *being-in-the-world* [22]. Construction involves the imposition of "new categories" that are not so-called natural categories in the Aristotelian sense but rather, categories that are based on features that guide retrieval. Communication is negotiated through an alignment of "external tokens" (ontic tags) associated with categories (ontological conceptualizations). There are no "pregiven conventions" or constraints to the communication of categories. "Communication is crucial, because it is the motor for testing the concepts' adequacy and for pushing the development of new concepts when there are misunderstandings of task failures" [38].

Interpretation results from the mutual adjustment of the explicit and implicit content of an utterance. An exhaustive, one-to-one mapping between concepts and words is quite implausible. An interpretation that does not match exactly the intent is not a failure of communication, rather "an illusion of the code theory that communication aims at the duplication of meanings" [37]. Communication succeeds despite semantic discrepancies because the words used in a given situation points the hearer in the direction intended by the speaker. It does not matter whether or not a word linguistically encodes a full-fledged concept, and, if so, whether it encodes the same concept for both speaker and hearer. Words are used as pointers to contextually intended senses; utterances are merely pieces of evidence of the speaker's intention. We need to know who the speaker is, their identity, in order to interpret the perspective from which the tag originates. The fact that the interpretation of tags is not exact reflects the real-world experience of communication and knowledge sharing and the need for an interactive, hermeneutic discourse to achieve understanding.

The ECM mandala captures where this communication takes place and where the alignment of meaning happens, not how. The model assumes an interpretive process of meaning making—a Gadamerian hermeneutic where a fusion of horizons occurs [16]. It is represented in the model as the circular yellow "cloud" in the overlapping schema clouds. This is the interactive space where ontological conceptualizations are transformed into ontic signs as tags, where intrapersonal schemas become cultural schemas. The space in which semantics emerge is the same space in which culture emerges.

5   MODIFIED SEMIOTICS

Tagging is a simple concept and a simple process: allow users to categorize entities using words that reflect the conceptual connections they make to the entity, and by doing so create a set of keywords that facilitates recall. Allowing users to create tags that are meaningful to them avoids the need for users to learn complex taxonomies created by experts, assuming experts could categorize the increasing amount of information being made available online. The tagging concept is based primarily on semiotic theory. Users associate signs (signifiers) with things signified (see Figure 3).
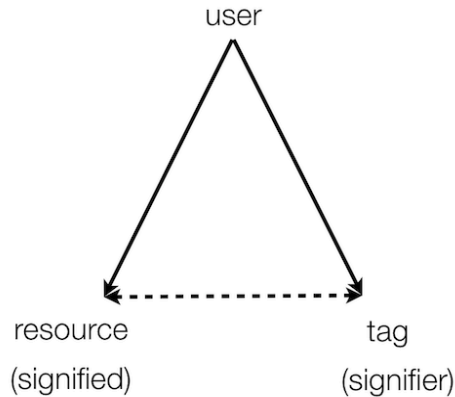
Figure 3. Simplified Semiotic Diagram

Semiotic approaches assume a direct connection between the sign and that which is signified—in our case the tag and the entity. Their focus is on the extrapersonal structures, and they tend to blackbox the intrapersonal schemas. However, as pointed out above, this assumption fundamentally confuses the social and cultural dimensions of tags [32]. The ECM conceptualizes the semiotic relationships differently. There is no direct or assumed connection between the tag and the entity. Rather, the interaction between the user and the entity is the emergent space in which the tag and its semantics are created. When a user interacts with his own tag as a consumer, the interaction between the user and the lexical unit generates a memory of the emergent experience with the entity. We can't say the same is true when a user interacts with a tag created by someone else. The chances that a similar semantic experience will occur is greater if the users share a cultural identity, but not if they don't. Assuming a relationship between the signifier and signified across users is tenuous and is avoided in the ECM. The semiotic relationship in the ECM can be illustrated as follows:
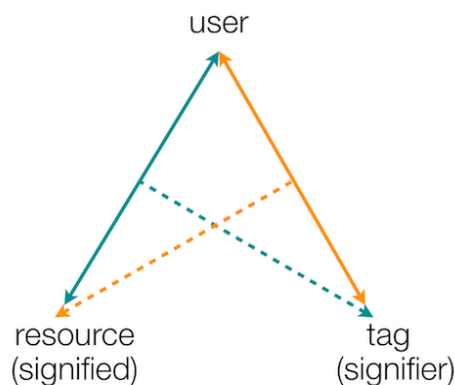

Figure 4. ECM Approach to Semiotic Relationships

Figure 4, depicts the interactions of the user as solid lines between the user and resource and the user and tag. In the user-resource interaction, a tag is generated through the experience, illustrated as a dashed line. The tag doesn't directly interact with the resource and isn't necessarily a direct referent for the resource, but rather links to a salient entry point of the schematic

associations that emerge as part of the cultural interaction between user and resource. In the user-tag interaction, the memory of the experience with the resource is generated (again, indicated with a dashed line). The interactions with extrapersonal structures are depicted as solid lines, while the activation of intrapersonal schemas are depicted with dashed lines. The semantics of the experience cannot be separated from the user and the cultural schemas that he uses during his interaction, as they would be in Figure 3 where the dashed line is indicated. Modeling the semiotic relationships with the ECM approach maintains the basic understanding of semantics as emerging through the interplay of extrapersonal structures (entities, tags) and intrapersonal schemas (user).

This variation on the semiotic model requires us to consider the cultural dimensions of tags, not just the social ones. In order to be used effectively to discern the semantics of tag sets, it requires us to identify intrapersonal schemas and focus on the interactive qualities of the tagging process. The biggest challenge in using this model to structuring folksonomies, in assessing their emergent semantics, is: how do we gather social metadata tags that reflect the various cultural schemas and identities of the users that the model suggests are required to properly discern the emergent semantics of folksonomies? How can we know what cultural identity a person is using at the particular time he is creating a tag? There is no simple answer to these questions, and the ECM does not specify a means by which to gather such information. It may be possible to have users create tags for their own cultural identities (e.g., male, student, New Englander, photographer, technologist, socialist, gay, atheist). But this, in some ways, reproduces the category problem discussed above (i.e., `DifferentFrom` and `SameAs`). It may be possible to mine the data from existing sources and create the cultural identity tags based on a variety of demographic information. Such information must be created, collected, and utilized if we are to properly disambiguate folksonomies to discern their emergent semantics. The next section discusses how we can structure the relationship between entity tags and identity tags once we have them.

6   CULTURAL IDENTITY TAG LAYERS
The relative simplicity of the tagging concept is transformed into a problem of greater complexity when we begin aggregating tags into tagclouds and broad folksonomies associated with particular perspectives—cultural identities and schemas. Compounding this complexity is the fact that many perspectives exist as part of an individual's cognition, and that the same perspective can be used as an identity for many individuals. To illustrate this complexity, I have extended the graph model for an e-commerce recommender system [23] from a 2-layer to a 4-layer graph. (The reader might imagine the ECM simplified even further into objectified layers of tags and rotated ninety degrees clockwise.) The detailed graph in Figure 5 depicts a relationship (from bottom to top) between an entity, the tags created for it by an individual, the individual, and the individual's multiple cultural identities.
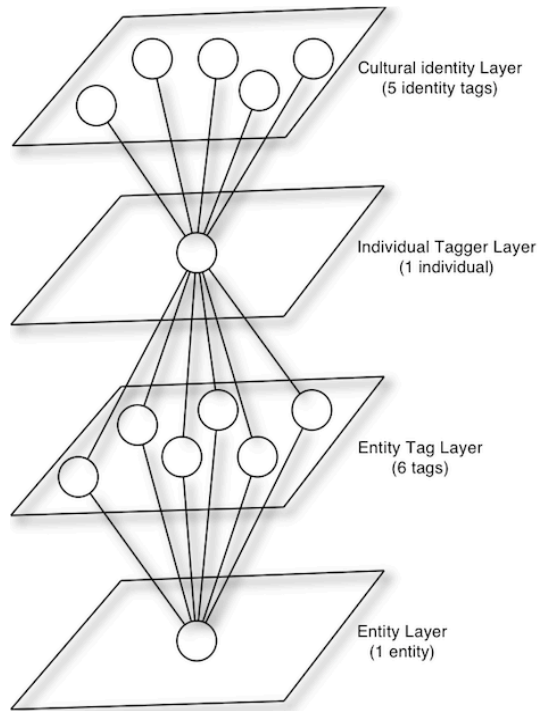
Figure 5. Four-layer graph model linking cultural identity of a single individual with a single entity.

This graph is an intermediate step. It still has the problem of not being able to associate particular entity tags with particular cultural identity tags. It represents the collective approach to examining folksonomies such that all individuals are aggregated into a collective supra individual for the purposes of analyzing the emergent semantics of tags. What we need, according to the ECM, is an increase in the number of individual taggers, which reflects the social nature of the metadata.

The graph in Figure 6 illustrates a small increase in this complexity by incorporating three individual taggers creating tags for the single entity. Once the cultural identities have been associated with the taggers, "activating" one of the identities reveals the individuals and the entity tags associated with that identity. We see where there is overlap between the tags (blue), indicating a cultural association for the identity, and where they diverge (red), indicating a more idiosyncratic association. Figures 5 and 6 illustrate the complexity of our task of associating semantic tags with phenomena and entities, generated by individuals using a variety of cultural identities with which we can then use to appropriately interpret the emergent semantics of tag sets. These graphs are small-scale and limited, and they do not represent the emergent nature of the ECM. They are only presented here as rough approximations of the associations that can be made using the ECM.
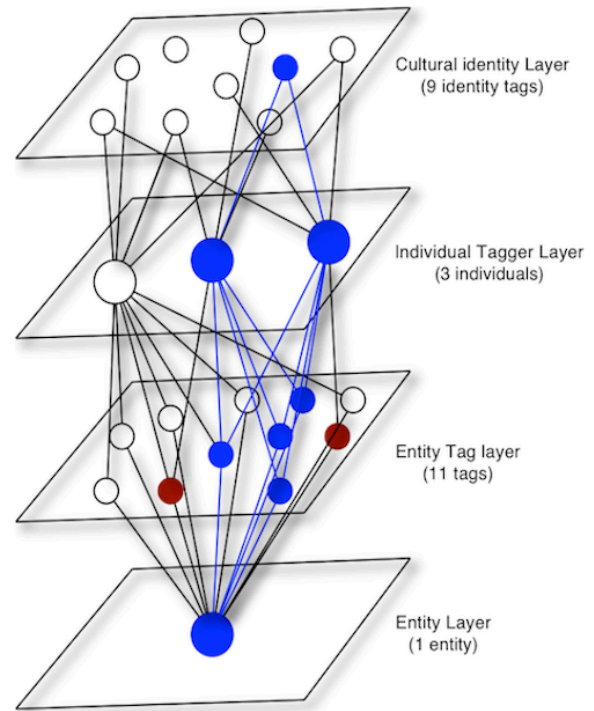

Figure 6. Activating a cultural identity shared by two individuals reveals the overlap of four of six tags created for the phenomenon.

7 CONCLUDING REMARKS

If we intended to capitalize on the tagging phenomenon, and the vast amount of social metadata available to us in order to discern an emergent semantics using folksonomies, we need to structure properly the relationships between entity, tag, and user as part of an interactive and emergent phenomenon of culture. I have presented here an emergent model of culture that can be used to help discern these emergent tag semantics. The model arises from the need to include the cultural dimensions of tags rather than rely solely on an analysis of their social dimensions. The ECM considers the emergent nature of culture and the hermeneutic way in which semantics are created within it. It embeds a refinement of the traditional semiotic approach to tag analysis. The ECM gives us a structure to discern the emergent semantics within folksonomies and through the incorporation of tags as lexical extrapersonal structures enables us to connect to representational formal ontologies [33]. Still, more work needs to be done: 1) with respect to how to identify sets or clusters of schemas that roughly correspond to cultural identities, and 2) exploring how we might extend the interactivity of the ECM and less-structured folksonomies that emerge from enacted experience to more strongly-structured computational ontologies.

8   REFERENCES

[1]   Anderson, R.C., Spiro, R.J. and Montague, W.E. (eds.). *Schooling and the acquisition of knowledge*. Lawrence Erlbaum, Hillsdale, NJ, 1984.

[2]   Aristotle, 1941 trans. *The Basic Works of Aristotle*. Random House, New York.

[3]   Aurnhammer, M., Hanappe, P. and Steels, L., 2006, Integrating collaborative tagging and emergent semantics for image retrieval. Paper presented at Proceedings of Collaborative Web Tagging Workshop, WWW2006, Edinburgh, Scotland, UK, May 22-26, 2006.

[4]   Bishr, M. and Kuhn, W. Geospatial Information Bottom-Up: A Matter of Trust and Semantics. in *The European Information Society*, 2007, 365-387.

[5]   Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V., Loreto, V., Hotho, A., Grahl, M. and Stumme, G., 2007, Network properties of folksonomies. Paper presented at WWW2007, Banff, Canada, May 8-12, 2007.

[6]   Chandrasekaran, B., Josephson, J.R. and Benjamins, V.R. 1999. What are ontologies, and why do we need them? *Intelligent Systems and Their Applications, IEEE [see also IEEE Intelligent Systems]*, *14* (1). 20-26.

[7]   Chiao, J.Y., Harada, T., Komeda, H., Li, Z., Mano, Y., Saito, D., Parrish, T.B., Sadato, N. and Iidaka, T. 2010. Dynamic Cultural Influences on Neural Representations of the Self. *Journal of Cognitive Neuroscience*, *22* (1). 1-11. doi:10.1162/jocn.2009.21192.

[8]   Cudré-Mauroux, P., Aberer, K., Abdelmoty, A., Catarci, T., Damiani, E., Illaramendi, A., Jarrar, M., Meersman, R., Neuhold, E., Parent, C., Sattler, K.-U., Scannapieco, M., Spaccapietra, S., Spyns, P. and De Tré, G. Viewpoints on Emergent Semantics. in *Journal on Data Semantics VI*, 2006, 1-27.

[9]   D'Andrade, R., 1995. *The Development of Cognitive Anthropology*. Cambridge University Press, Cambridge.

[10] D'Andrade, R.G. and Posner, M.I. Cultural Cognition. in Posner, M. ed. *Foundations of Cognitive Science*, The MIT Press, Cambridge, MA, 1989, 795-830.

[11] Davis, P.M. Cognition and learning: A review of the literature with reference to ethnolinguistic minorities, Summer Institute of Linguistics, Dallas, TX, 1991, 80.

[12] DiMaggio, P. 1997. Culture and cognition. *Annual Review of Sociology*, *23*. 263-288.

[13] Fauconnier, G. and Turner, M. 1998. Conceptual Integration Networks. *Cognitive Science*, *22* (2). 133-187.

[14] Fisher, K. Locating Frames in the Discursive Universe *Sociological Research Online, 2(3)*, 1997.

[15] Fonseca, F. 2007. The double role of ontologies in information science research. *Journal of the American Society for Information Science and Technology*, *56* (6). 786-793.

[16] Gadamer, H.-G., 1975. *Truth and Method*. Continuum, New York.

[17] Golder, S. and Huberman, B. 2006. The Structure of Collaborative Tagging Systems. *Journal of Information Science*, *32* (2). 198-208.

[18] Gruber, T., 2005, Ontology of Folksonomy: A Mash-up of Apples and Oranges. Paper presented at First on-Line conference on Metadata and Semantics Research (MTSR'05), http://www.metadata-semantics.org/.

[19] Gruber, T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. in Guarino, N. and Poli, R. eds. *Formal Ontology Conceptual Analysis and Knowledge Representation*, Kluwer Academic Publishers, 1993.

[20] Hak Lae, K., Passant, A., Breslin, J.G., Scerri, S. and Decker, S., 2008), Review and Alignment of Tag Ontologies for Semantically-Linked Data in Collaborative Tagging Spaces. in *Semantic Computing, 2008 IEEE International Conference on*, (*4-7 Aug. 2008*, 315-322.

[21] Halpin, H., Robu, V. and Shepherd, H., 2007), The complex dynamics of collaborative tagging. in *WWW '07: Proceedings of the 16th international conference on World Wide Web*, ACM Press, 211-220.

[22] Heidegger, M., 1927. *Being and Time*. Harper and Row, New York.

[23] Huang, Z., Chung, W. and Chen, H. 2004. A graph model for E-commerce recommender systems. *Journal of the American Society for Information Science and Technology*, *55* (3). 259.

[24] Hutchins, E., 1995. *Cognition in the Wild*. The MIT Press, Cambridge, MA.

[25] Hutchins, E., 2000, The appropriation of cultural structure by cognitive processes. Paper presented at American Anthropological Association, San Francisco, CA.

[26] Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-network-theory*. Oxford University Press.

[27] Latour, B., 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press.

[28] Mazzocchi, S., 2005, Folksologies: de-idealizing ontologies, Stefano's Linotype, Accessed: 2008.05.17.

[29] Moscovici, S. The Phenomenon of Social Representations. in Farr, R.M. and Moscovici, S. eds. *Social Representations*, Cambridge University Press, London, 1984.

[30] Rumelhart, D.E. and McClelland, J.L. Parallel Distributed Processing: Exploration in the microstructure of cognition, Vols. 1 & 2. in *Psychological and Biological Models*, The MIT Press, Cambridge, 1986.

[31] Saab, D.J., 2009), Culture as Mediator for what is *Ready-to-hand*: A Phenomenologica Exploration of Semantic Networks. in *North American Conference on Computing and Philosophy, NA-CAP '09*, (*Indiana University, Bloomington, Indiana, USA, June 14-17, 2009*.

[32] Saab, D.J., 2010), The Ontology of Tags. in *Proceedings of the iConference 2010*, (*Urbana-Champaign, IL, February 3-6, 2010*.

[33] Saab, D.J. and Fonseca, F., 2008), Ontological Complexity and Human Culture. in *Philosophy's Relevance in Information Science*, (*Paderborn, Germany, October 3-4, 2008*.

[34] Saab, D.J. and Riss, U.V. Logic and Abstraction as Capabilities of the Mind: Reconceptualizations of Computational Approaches to the Mind (to appear). in Vallverdú, J. ed. *Thinking Machines and the Philosophy of Computer Science: Concepts and Principles*, IGI Global, Hershey, PA, 2009.

[35] Schmitz, C., Hotho, A., Jäschke, R. and Stumme, G., 2006), Mining Association Rules in Folksonomies. in *Data Science and Classification. Proceedings of the 10th IFCS Conf.*, (*Heidelberg, July 2006*), Springer, 261-270.

[36] Smith, B. Ontology. in Floridi, L. ed. *Blackwell's Guide to Philosophy of Computing and Information*, Blackwell, Oxford, 2003, 155-166.

[37] Sperber, D. and Wilson, D. The mapping between mental and public lexicon. in Carruthers, P. and Boucher, J. eds.

Thought and Language, Cambridge University Press, Cambridge, 1998.

[38] Staab, S., Santini, S., Nack, F., Steels, L. and Maedche, A. 2002. Emergent semantics. *IEEE Intelligent Systems*, *17* (1). 78-86.

[39] Stillings, N.A., Weisler, S.W., Chase, C.H., Feinstein, M.H., Garfield, J.L. and Rissland, E.L., 1995. *Cognitive Science: An Introduction*. The MIT Press, Cambridge, MA.

[40] Strauss, C. and Quinn, N., 1997. *A cognitive theory of cultural meaning*. Cambridge University Press, Cambridge.

[41] Talmy, L. 2001. The cognitive culture system. *Monist*, *78* (1).

[42] Tanaka, J. and Taylor, M. 1991. Object Categories and Expertise: Is the Basic Level in the Eye of the Beholder? *Cognitive Psychology*, *23* (3). 457-482.

[43] Vander Wal, T. Online Information Folksonomy: Presentation Posted *Personal InfoCloud*, *personalinfocloud.com*, personalinfocloud.com, 2006.