Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge

Dominik Benz
Knowledge and Data Engineering Group (KDE)
University of Kassel, Germany
benz@cs.uni-kassel.de

Andreas Hotho
Department for Artificial Intelligence and Applied Computer Science
University of Würzburg, Germany
hotho@informatik.uni-wuerzburg.de

Stefan Stützer
Knowledge and Data Engineering Group (KDE)
University of Kassel, Germany
stuetzer@cs.uni-kassel.de

Gerd Stumme
Knowledge and Data Engineering Group (KDE)
University of Kassel, Germany
stumme@cs.uni-kassel.de

Keywords

ontology learning, web2.0, semantics, synsetized folksonomy

ABSTRACT

The participatory nature of many Web 2.0 platforms makes a large portion of users' interactions with each other and with information resources digitally observable. The assumption that the evolving structure of these digital records contains implicit evidences for the underlying semantics has been proven by successful approaches of making the emergent semantics explicit, e.g. in the form of lightweight ontologies.

In this paper, we provide further evidence for the great potential of self-emerging ontologies from Web 2.0 data, exemplified by collaborative tagging systems. We hereby combine and extend prior research, where we identified crucial aspects for successful methods to infer tag semantics. The additional contribution of this paper is to propose an extended methodology to induce a hierarchical organization scheme from the initially flat tag space which captures the semantics and the diversity of the shared knowledge. It comprises the introduction of a *synsetized folksonomy* (which tackles the problem of synonymous tags) and a clustering approach for tag sense disambiguation.

In order to assess the quality of the learned semantics, we compare the inferred organization scheme with manually built categorization schemes from WordNet and Wikipedia. Our results exhibit clear similarities; so in summary, our work demonstrates a successful example of self-emergent ontologies from Web 2.0 data.

1. INTRODUCTION

The prototype theory of semantics proposes that systems of categories are rooted in people's experience of interacting with the world and each other. This implies that concepts are not simply "out there", but are learned and constructed in an interactive manner. The participatory nature of many Web 2.0 platforms transports a large amount of such interactions into public space, where they become available as digital records in the form of e.g. collaboratively tagged content, Twitter feeds, social networks, weblogs and many others. The obvious assumption that the evolving structure of this kind of data contains implicit evidences for the underlying semantics has been proven by successful approaches of making the emergent semantics explicit, e.g. in the form of lightweight ontologies.

In this paper, we provide further evidende for the great potential of self-emerging ontologies from Web 2.0 data, exemplified by collaborative tagging systems. We hereby combine and extend two branches of our prior research, where crucial aspects for successful methods to harvest semantics have been identified:

- An in-depth understanding of the characteristics of similarity measures between objects in a folksonomy. We demonstrated a semantic grounding procedure for tag and resource similarity [6, 22] and identified e.g. measures to detect synonym relationships among tags.
- An algorithm to induce a hierarchical organization scheme from the initially flat tag space [4]. Its main building blocks comprise measures of of tag relatedness and tag generality.

Based on this experience, the additional contribution of this paper is to extend our originally proposed algorithm, informed by the results of our work on semantic grounding of tag relatedness. By doing so, we enhance the algorithm such that it creates a hierarchical organization scheme which captures the semantics and the diversity of the shared knowledge.

By "diversity", we mean that this self-emergent ontology integrates different views on the data: Naturally, an emerging concept can be described by a variety of tags. On the other side, the same tags might be used in several communities, but with a different meaning. We account for both aspects by first grouping together tags with a very similar meaning (resulting in a so-called *synsetized folksonomy*), and then disambiguating the different senses while assembling the hierarchical relationships.

Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.
.

In order to assess the quality of the learned semantics, we compare the inferred organization scheme with manually built categorization schemes from WordNet and Wikipedia. Our results exhibit clear similarities; so in summary, our work demonstrates a successful example of self-emergent ontologies from Web 2.0 data.

The rest of this paper is structured as follows: We start with a description of relevant related work in section 2. As we propose that an appropriate preprocessing (namely the idenfication of synonym and ambiguous tags) of the "raw" folksonomy data has beneficial effects for the applicattion of an ontology learning algorithm, we detail in the following Section 3 on each preprocessing step. Section 4 presents our proposed algorithm to induce hierarchical relationships among tags. Section 5 introduces an experimental methodology to evaluate the learned ontology, before we discuss the results in Section 6. We finally conclude an point to possible future research directions and implications in Section 7.

2. RELATED WORK

The origins of automatic acquisition of semantics from unstructured and semi-structured resources can be seen can be found in ontology learning from text [19]. Ref. [9] provides an approach to synonym acquisition from text corpora.

With the advent of the Web 2.0 paradigm, early studies like Ref. [23] introduced several concepts of bottom-up social annotation and stimulated the discussion of emergent semantics within this kind of data. Ref. [13, 18, 23] provide overviews of the strengths and weaknesses of such systems. Ref. [17, 24] introduce a tri-partite graph representation for folksonomies, where nodes are users, tags and resources. Ref. [11] provides a first quantitative analysis of del.icio.us. We investigated the distribution of tag co-occurrence frequencies in Ref. [7] and the network structure of folksonomies in Ref. [8].

Recent work like Ref. [6, 22] provided a characterization of several measures of tag and resource relatedness and identified e. g., measures able to detect synonym tags. This work is an example of a considerable number of investigations which aremotivated by the vision of "bridging the gap" between the Semantic Web and Web 2.0. These works can broadly be divided into two classes of approaches - the first class tries to infer semantics from the Web 2.0 data alone, without the consultation of existing semantic repositories like ontologies. As an example, Ref. [24] provides a model of semantic-social networks for extracting lightweight ontologies from del.icio.us. Other approaches for learning taxonomic relations from tags are provided by Ref. [14, 30]. Ref. [12] presents a generative model for folksonomies and also addresses the learning of taxonomic relations. Ref. [31] applies statistical methods to infer global semantics from a folksonomy.

The second class of approaches tries to establish a connection between tags and external semantic resources, yielding a semantic enrichment of the folksonomy tag space [1]. Ref. [21] proposes the integration of expert-created and user-generated metadata in the form of so-called collabularies. On a more general level, Ref. [20] discusses several possible contexts for the disambiguation of tag senses.

All ontology learning approaches face the difficult problem of assessing the quality of the learned semantics; Ref. [10, 2] analyze possible methodologies and similarity measures for gold-standard based evaluation approaches.

To put the current work in this context, this paper mainly combines two mentioned branches of research — more precisely, we apply the insights about the characteristics of tag relatedness measures presented in [6] to enhance the ontology learning algorithm presented in [4].

3. FOLKSONOMY PREPROCESSING

Most existing approaches to infer semantic relationships among tags in a folksonomy work "directly" on the containted tags and the underlying folksonomy network. This intuitive approach neglects two fundamental properties of tags — namely that (i) naturally several tags are used to describe the same semantic concept (we refer to this as *synonym tags*[1]) and that (ii) a single tag can have several meanings (we will denote this as an *ambiguous tag*.)

We argue that it is resonable to address both factors in a *preprocessing step*: First, synonym tags are identified and grouped together; and second, for each tag in the resulting "synsetized folksonomy" the number of its meanings is computed. Both processes are described in detail in the following two subsections; we start with a definition of the folksonomy model we use.

3.1 Folksonomy model

In the following we will use the definition of folksonomy provided in [16]:

Definition A *folksonomy* is a tuple $\mathbb{F} := (U, T, R, Y)$ where U, T, and R are finite sets, whose elements are called *users*, *tags* and *resources*, respectively. Y is a ternary relation between them, i. e., $Y \subseteq U \times T \times R$. A *post* is a triple (u, T_{ur}, r) with $u \in U$, $r \in R$, and a non-empty set $T_{ur} := \{t \in T \mid (u, t, r) \in Y\}$.

3.2 Dataset

Because the following explanations are understood best when explained with real-world examples, we detail now on the dataset used for our experiments. It was crawled in July 2005 from the social bookmarking site del.icio.us[2]. Originally, it contained $|U| =$ 75.260 users, $|T| = 533.191$ tags, and $|R| = 3.151.353$ resources, related by $|Y| = 17.364.552$ triples. As most of our following work is based on the *co-occurrence* between tags, we removed all tags used only once by a single user; this left us with a dataset of $|U| = 74.680$ users, $|T| = 373.690$ tags, and $|R| = 2.972.695$ resources, related by $|Y| = 17.181.896$ triples.

3.3 Synonym Identification

As stated above, the first task we wish to accomplish is to detect synonym tags. More precisely, we want to "shrink" the folksonomy's vocabulary T by merging all tags with a very similar meaning into a single new "artificial" tag. As an example, consider that the tags $t_1, t_2 \in T$, $t_1 = car$ and $t_2 = automobile$ are present. They both refer to the same semantic concept of a wheeled motor vehicle. Then we would create a *synset*[3] $s_1 = \{t_1, t_2\}$ and replace each occurrence of t_1 and t_2 in the folksonomy by s_1.

As we replace in this step each tag $t \in T$ by a synset (which contains at least t itself, and possibly other tags which have the same meaning as t), we will refer to the resulting structure as a *synsetized folksonomy*. More formally, we define the following:

Definition A *synsetized folksonomy* is a tuple $\mathbb{F}^s := (U, S, R, Y^S)$ where U and R are the sets of users and resources present in a folksonomy $\mathbb{F}$. S is the synsetized vocabulary of F, whereby each original tag $t \in T$ has been replaced by its containing synset $s \subset S$.

[1] We are aware that this usage of the term *synonym* is not precise in a linguistical sense; however, for the topic of this paper it is sufficient to refer to two tags as being synonyms when they refer to the same semantic concept.

[2] http://www.delicious.com

[3] We use the term "synset" in the same way it is used in WordNet [25] — a synset is a set of words with a semantically equivalent meaning.

Synset	number of users
tool tools	14938
computer computers	8285
color colors colour	6094
code coding	5821
folksonomies folksonomy	3080
objectoriented oo oop	1500
a2004 a2005 ambient band bands beatles bootleg bootlegs coldplay concerts dj drums electro electronic electronica folk guitar harmonica hiphop idm ilm indie instrument instruments jrock label labels lyrics m2005 mashup mashups metal mixes mp3 mp3s mu...	15124

Table 1: Example synsets created by a similarity threshold of *min_syn* $= 0,96$.

The synsets $s \in S$ may overlap and contain exhaustively all tags $t \in T$, i. e., $\bigcup_{s \in S} = T$. Each original tag t is mapped to its synset by the synset function $syn : T \rightarrow \mathbb{P}(T)$. In other words, t is mapped to a subset of T.

The synsetized tag assignments Y^S are created by replacing each triple $(t, u, r) \in Y$ with $(syn(t), u, r)$.

It is important to notice that the synsetized vocabulary is at most as big as the original vocabulary, i. e., $|S| \leq |T|$ and $|Y^S| \leq |Y|$. We argue that a properly established synsetized folksonomy is able to overcome the synonymy problem inherent in collaborative tagging systems. Of course, the crucial point here is to define the mapping function syn. An important input when defining such a function are measures of semantic similarity among tags. Given such a valid measure, an intuitive solution is to group together all tags with a very high degree of semantic similarity above a given treshold.

In [6], we examined several measures of tag relatedness and identified the *tag context relatedness* (denoted as *TagCont*) as the best means to discover synonym tags within a folksonomy.

TagCont is computed in the vector space $\mathbb{R}^T$, where, for tag t, the entries of the vector $\vec{v}_t \in \mathbb{R}^T$ are defined by $v_{tt'} := w(t, t')$ for $t \neq t' \in T$, where w is the co-occurrence count[4].

Similarity is then determined by using the cosine measure, a measure customary in Information Retrieval [28]: If two tags t_1 and t_2 are represented by $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^T$, their cosine similarity is defined as: $\text{cossim}(t_1, t_2) := \cos \measuredangle(\vec{v}_1, \vec{v}_2) = \frac{\vec{v}_1 \cdot \vec{v}_2}{||\vec{v}_1||_2 \cdot ||\vec{v}_2||_2}$.

In order to create the synsets, we followed the intuitive approach described above and computed the pairwise tag context similarity among all tags t in our folksonomy. Then, we grouped together all tags whose similarity was above a given threshold tau_{sim}. More precisely, we defined the mapping function syn as follows (sim corresponds to the tag context similarity described above):

$$syn(t) = \{t' \in T : sim(t, t') > \tau_{sim}\}$$

Please note that because $sim(t, t) = 1$ a tag t is always contained in its "own" synset, i. e., $\forall t \in T : t \in syn(t)$. In our experiment, it turned out that a similarity threshold of $\tau_{sim} = 0.96$ gave the best results. Lowering the treshold turned out to group together relatively unrelated tags, while putting it higher resulted in a very small number of grouped tags.

After computing all synsets, the synsetized folksonomy contained $|S| = 373.572$ synsets and $|Y^s| = 17.154.948$ synsetized tag assignments. Table 1 shows some exemplary synsets created by this approach. One can see that obviously some tags have been grouped in a rather semantically precise way (e. g., *objectoriented oo oop*), while there also exist relatively large synsets with broadly related terms (last row of Table 1). However, in general the quality of the obtained synsets can be regarded as satisfactory.

[4]We set $v_{tt} = 0$ because we want two tags to be considered related when they occur in a similar context, and not when they occur together; see [6] for details.

tag	preference tags
language	english words linguistics dictionary writing education learning
language	java dev code
paper	origami crafts
paper	research article read articles writing papers science diy
training	education tutorial tutorials tips resource article learning webdesign
training	fitness running
ai	games game
ai	science intelligence research learning ia robots java algorithms
ipod	apple mac
ipod	gadgets shopping
ipod	music podcast podcasting mp3 audio itunes
nyc	food
nyc	music
nyc	newyork events travel urban photography photo shopping
nyc	politics

Table 2: Excerpt of discovered ambiguous tags along with their preference tags, using a similarity threshold of $min\_sim = 0,55$.

3.4 Discovering ambiguous tags

In the previous section, a major problem of collaborative tagging systems has been addressed - namely that different people use different terms for the same semantic concept. Our proposed *synsetized folksonomy* helps to tackle this problem. Apart from that, another common issue is that a single term can have several meanings - this phenomenon is commonly referred to as *homonymy*. Hence another crucial task when trying to infer ontologies from folksonomies is to identify ambiguous tags.

Based on the obtained synsetized folksonomy from the previous step, our next step is to iterate over all synsets (represented by an artificially introduced tagid) and to check whether the current synset has several meanings. This check is performed in two steps:

1. *Context identification:* We first identify the context of a given tag. We assume that the different meanings of an ambiguous tag t are reflected well within the other tags t was used together with. In other words, our hypothesis is that the co-occurring tags of t provide a valid context to discover different meanings of t. In our study, we used the $n = 10$ most frequently co-occuring tags as a context.

2. *Context disambiguation:* Based on the most frequently co-occuring tags, our goal is to discover the different meanings present in this context. To this end we once again rely on the tag context relatedness described above. We first represent each co-occuring tag t_{cooc} by its tag context vector. We then apply a standard average-link hierarchical clustering algorithm [26] based on the cosine similarity between the obtained vectors. The actual clusters are obtained by applying a simlarity threshold min_sim.

This two-step approach allows us to identify the different meanings of ambiguous tags. As we derive the different senses from the co-occuring tags, we have the possibility to label each resulting sense with the co-occuring tags representing the resprective sense. We refer to the latter as *preference tags*.

Table 2 shows an example of ambiguous tags created by this procedure with a similarity threshold of $min\_sim = 0.55$

Taking the tag *paper* as an example, one can see that our approach clearly identifies two different meanings - on the one hand scientific papers (described by the preference tags *research article read articles writing papers science diy*), and on the other hand a sense related to handicrafts (described by *origami crafts*). Another example is the tag *language*, which is correctly disambiguated into a linguistic and a technical sense.

After this step, we have identified synonymous and homonymous tags. This information will now serve as an input to an algorithm to infer hierarchical relationships among tags.

4. ONTOLOGY LEARNING APPROACH

The goal of the algorithm described now is to automatically induce a concept hierarchy, i.e., a tree structure, whose nodes (representing concepts) each consist of one or more tags from a folksonomy. Concept specificity increases with increasing depth in the tree, and there exists only a single type of relation, whose semantics resembles closely the one of the taxonomic relation. Before detailing on our proposed algorithm, we briefly given an overview of classes of approaches to reach this goal.

4.1 Classes of approaches

Existing approaches to infer hierarchical tag relationships can broadly be assigned to one of the following three classes:

- *Social Network Analysis:* [24] pioneered in applying centrality and other measures like the clustering coefficient coming from social network analysis to the tag co-occurrence networks in order to identify broader and narrower terms. [14] proposed betweeness centrality as tag generality measure. The latter approach is the basis of [5].
- *Statistical approaches:* The work of [30] and [29] is based on statistical models of tag subsumption, the latter is corroborated with the theory of association rule mining.
- *Clustering approaches:* Starting from a similarity measure between tags, clustering approaches like [3] identify groups of highly related tags. Depending on the chosen clustering algorithm, a hierarchical relationship between the tag clusters is established.

4.2 Proposed Algorithm

The proposed algorithm is an extension of the work of [4], which itself is an extension of [14]. It comprises the following steps:

1. Filter the tags by an occurrence threshold τ_{occ}
2. Order the tags in descending order by generality (measured by degree centrality [15] in the tag-tag co-occurrence network)
3. Starting from the most general tag, add all tags t_i subsequently to an evolving tree structure.
   (a) identify the most similar existing tag t_{sim} to t_i (using the co-occurrence weights as similarity measure). If t_i is very general (determined by a generality threshold min_gen) or no sufficiently similar tag exists (determined by a similarity threshold min_sim), append t_i underneath the root node of the hierarchy.
   (b) If t_{sim} is an ambiguous tag, identify the correct sense of t_{sim} in the context of the t_i by taking into account the preference tags of t_{sim} and t_i
   (c) append t_i as a less general term underneath the correct sende of t_{sim}.
   (d) if t_i is an ambiguous tag, repeat steps 3.a - 3.c for each of its senses.
4. Apply a post-processing to the resulting tree by re-inserting orphaned tags underneath the root node in order to create a balanced representation. The re-insertion is done based on steps 3.a-3.c.

Compared to the original algorithm proposed in [4], our proposed algorithm provides a handling of ambiguous tags and the post-processing step. Both extensions lead to a clearer learned ontology and a more balanced hierarchical structure, which comes more closely to manually built hiearchies (as will be shown in the following chapter). Another extension is that we apply this algorithm to a synsetized folksonomy; because of that, the synonymy problem of collaborative tagging systems is addressed as well. In summary, this procedure yields a hierarchical organization scheme which integrates the “diverse” views on the data.

The control flow of the algorithm is influenced by a set of parameters (similarity threshold min_sim, generality threshold min_gen). We performed empirical experiments to optimize each parameter; the following section describes the evaluation setting used to determine the degree of success for a single run.

5. EVALUATION

Choosing a gold-standard based evaluation paradigm, it is a nontrivial task to judge the similarity between a learned concept hierarchy and a reference hierarchy, especially regarding the absence of well-established and universally accepted evaluation measures. Hence we are facing two crucial questions - first, which gold-standard ontology to choose, and second which measure to use to compute the similarity between the learned and the gold-standard ontology. In the following, we describe the reference ontologies (derived from WordNet and Wikipedia) and the evaluation measures used in our experiments. The general idea is always that a learned ontology is better the more similar it is to a manually built one.

5.1 Wordnet

WordNet [25] is a structured lexical database of the English language. It contains roughly 203.000 terms grouped into 115.400 synsets. Among the synsets, several relations are defined; the most important one is the taxonomic relation. As a first gold-standard, we extracted the taxonomic hierarchy among synsets in WordNet.

5.2 Wikipedia

Wikipedia[5] is the world's largest collaboratively built online encyclopedia. Its articles can be assigned to one or more categories from a collaboratively built category system. The underlying taxonomy has been made available by researchers [27]. Despite the fact that (as is customary for Wikipedia) everybody is able to create and mofify these categories, the Wikipedia category taxonomy can be seen as a lightweigt ontology. This was the reasons why we chose

[5]http://www.wikipedia.org

it as the second gold-standard for our evaluation purposes. The obtained final ontology is also considerably larger than the WordNet ontology and contains roughly 2.4 million concepts.

5.3 Evaluation Measures

Having two gold-standard ontologies at hand (as described above), the next crucial question is how to judge the similarity between a learned and reference ontology. Despite the fact that finding single valid similarity score for two (possibly very large) hierarchical structures is a non-trivial task, we build on existing work to this end. Dellschaft et al. [10] proposes two measures, namely *taxonomc precision* and *taxonomic recall* for this purpose. The basic idea is hereby to find a concept present in both ontologies, and then to extract a characteristic excerpt (consisting e. g., from the sub- and super-concepts) from both ontologies. If both excerpts are very similar, then the ontologies itself are judged to be similar. The same idea underlies the measure of *taxonomic overlap* proposed by Mädche [19]. We will adapt both measures for our evaluation, namely the *taxonomic F-measure* (computed as the harmonic mean between taxonomic precision and recall) as well as the taxonomic overlap.

Based on these measures and both gold-standard ontologies, we first ran several experiments to optimize the parameters. The next section presents the results from the best parameter settings found.

6. RESULTS

Ultimatively, the question addressed in the evaluation of our approach is to which extent our proposed method is able to reproduce manually built ontologies. Figure 1 summarizes the results obtained for the best parameter settings.

For both reference ontologies, our proposed extended algorithm yields ontologies which resemble more closely the gold-standard. We attribute this to the fact that the extensions comprise the preprocessing of the folksonomy by identifying synsets, and by taking into account ambibuity of tags while assembling the hierarchical structure. In this way, we can reproduce better the diversity of shared knowledge which is also inherent in manually built ontologies like the one derived from WordNet or Wikipedia.

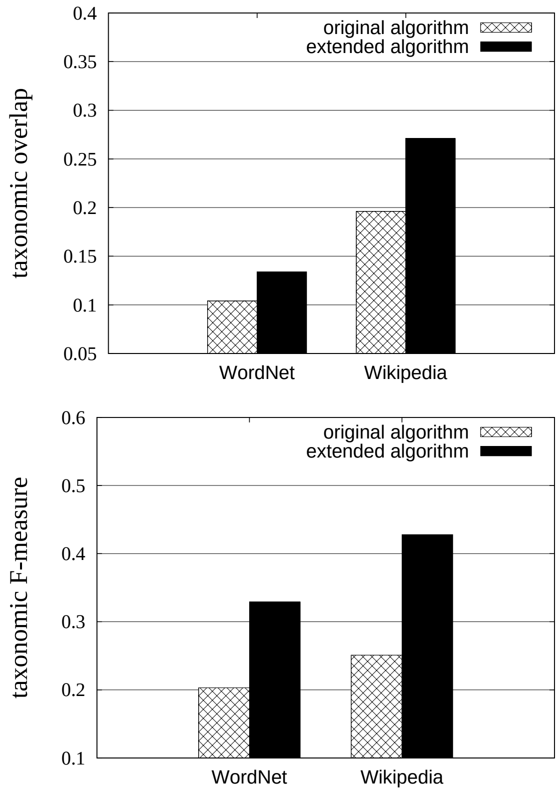
In order to give a visual impression of the results, Figures 2 and 3 depict excerpts of the learned ontology, extracted around the concept *programming* and *web*. The green nodes are synsetized tag nodes, while the yellow nodes signalize ambigous tags. Please note that e. g., the term *language* in the programming excerpt also found elsewhere in the ontology, representing its other (linguistic) sense. This is an example of a successful integration of the diverse views on the underlying data.

7. CONCLUSIONS

In this work we work demonstrated a successful example of self-emergent ontologies from Web 2.0 data. As explained at the beginning of the paper, the participatory nature of these systems makes a large amount of users' interactions with each other and with information resources digitally observable. Based on the example of collaborative tagging systems, we provided further evidence that the resulting structure of the digital records of such interactions contain implicit semantics.

We proposed an extended algorithm to induce a hierarchical organization scheme among the initially flat tag space within a folksonomy. As a crucial preprocessing step, we suggested the discovery of synonym and ambiguous tags, introducing the concept of a *synsetized folksonomy*. Along with an improved algorithmic scheme, these extensions lead to a learned ontology which resembles more closely a manually built gold-standard derived from WordNet and Wikipedia.

Figure 1: Experimental results of comparing the learned ontology with the reference ontologies from WordNet and Wikipedia. The synsetized folksonomies were crated with a synonymy threshold of $min\_syn = 0,96$.



For future work, we plan to integrate alternative measures to optimize the discovery of synonym and homonym tags. Intelligent tag normalization strategies could further improve the quality of the learned semantics. Furthermore, the evaluation of the learned ontologies in the context of an ontology-driven Semantic Web application might tackle the inherent shortcomings of our current gold-standard based evaluation scheme.

Acknowledgment

This research has been partly supported by the TAGora project (FP6-IST5-34721) funded by the Future and Emerging Technologies program (IST-FET) of the European Commission.

8. REFERENCES

[1] S. Angeletou. Semantic enrichment of folksonomy tagspaces. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 889–894. Springer, 2008.

[2] K. Bade and D. Benz. Evaluation strategies for learning algorithms of hierarchical structures. In *Proceedings of the 32nd Annual Conference of the German Classification Society - Advances in Data Analysis, Data Handling and*

Business Intelligence (GfKl 2008), Studies in Classification, Data Analysis, and Knowledge Organization, Berlin-Heidelberg, 2008. Springer. in press.

[3] G. Begelman, P. Keller, and F. Smadja. Automated tag clustering: Improving search and exploration in the tag space. In *Proceedings of the Collaborative Web Tagging Workshop at the WWW 2006*, Edinburgh, Scotland, May 2006.

[4] D. Benz and A. Hotho. Position paper: Ontology learning from folksonomies. In A. Hinneburg, editor, *Workshop Proceedings of Lernen - Wissensentdeckung - Adaptivität (LWA 2007)*, pages 109–112. Martin-Luther-Universität Halle-Wittenberg, Sept. 2007. http://lwa07.informatik.uni-halle.de/kdml07/kdml07.htm.

[5] D. Benz, K. Tso, and L. Schmidt-Thieme. Automatic bookmark classification: A collaborative approach. In *Proceedings of the Second Workshop on Innovations in Web Infrastructure (IWI 2006)*, Edinburgh, Scotland, 2006.

[6] C. Cattuto, D. Benz, A. Hotho, and G. Stumme. Semantic grounding of tag relatedness in social bookmarking systems. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *The Semantic Web – ISWC 2008, Proc.Intl. Semantic Web Conference 2008*, volume 5318 of *LNAI*, pages 615–631, Heidelberg, 2008. Springer.

[7] C. Cattuto, V. Loreto, and L. Pietronero. Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences (PNAS)*, 104:1461–1464, 2007.

[8] C. Cattuto, C. Schmitz, A. Baldassarri, V. D. P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. Network properties of folksonomies. *AI Communications Journal, Special Issue on Network Analysis in Natural Sciences and Engineering*, 20(4):245–262, 2007.

[9] P. Cimiano. *Ontology Learning and Population from Text — Algorithms, Evaluation and Applications*. Springer, Berlin–Heidelberg, Germany, 2006. Originally published as PhD Thesis, 2006, Universität Karlsruhe (TH), Karlsruhe, Germany.

[10] K. Dellschaft and S. Staab. On how to perform a gold standard based evaluation of ontology learning. In *Proceedings of ISWC-2006 International Semantic Web Conference*, Athens, GA, USA, November 2006. Springer, LNCS.

[11] S. Golder and B. A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, April 2006.

[12] H. Halpin, V. Robu, and H. Shepard. The dynamics and semantics of collaborative tagging. In *Proceedings of the 1st Semantic Authoring and Annotation Workshop (SAAW'06)*, 2006.

[13] T. Hammond, T. Hannay, B. Lund, and J. Scott. Social Bookmarking Tools (I): A General Review. *D-Lib Magazine*, 11(4), April 2005.

[14] P. Heymann and H. Garcia-Molina. Collaborative creation of communal hierarchical taxonomies in social tagging systems. Technical Report 2006-10, Computer Science Department, April 2006.

[15] B. Hoser, A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Semantic network analysis of ontologies. In *European Semantic Web Conference, Budva, Montenegro*, June 2006.

[16] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. In Y. Sure and J. Domingue, editors, *The Semantic Web: Research and Applications*, volume 4011 of *LNAI*, pages 411–426, Heidelberg, 2006. Springer.

[17] R. Lambiotte and M. Ausloos. Collaborative tagging as a tripartite network. *Lecture Notes in Computer Science*, 3993:1114, Dec 2005.

[18] B. Lund, T. Hammond, M. Flack, and T. Hannay. Social Bookmarking Tools (II): A Case Study - Connotea. *D-Lib Magazine*, 11(4), April 2005.

[19] A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academic Publishers, 2002.

[20] C. man Au Yeung, N. Gibbins, and N. Shadbolt. Contextualising tags in collaborative tagging systems. In *HT '09: Proceedings of the Twentieth ACM Conference on Hypertext and Hypermedia*, New York, NY, USA, July 2009. ACM.

[21] L. B. Marinho, K. Buza, and L. Schmidt-Thieme. Folksonomy-based collabulary learning. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 261–276. Springer, 2008.

[22] B. Markines, C. Cattuto, F. Menczer, D. Benz, A. Hotho, and G. Stumme. Evaluating similarity measures for emergent semantics of social tagging. In *18th International World Wide Web Conference*, pages 641–641, April 2009.

[23] A. Mathes. Folksonomies – Cooperative Classification and Communication Through Shared Metadata, December 2004. http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html.

[24] P. Mika. Ontologies are us: A unified model of social networks and semantics. In *International Semantic Web Conference*, LNCS, pages 522–536. Springer, 2005.

[25] G. A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38(1):39–41, 1995.

[26] P. Pantel and D. Lin. Document clustering with committees. In *In Proc. of SIGIR02*, pages 199–206. ACM Press, 2002.

[27] S. P. Ponzetto and M. Strube. Deriving a large-scale taxonomy from wikipedia. In *AAAI*, pages 1440–1445. AAAI Press, 2007.

[28] G. Salton. *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

[29] C. Schmitz, A. Hotho, R. Jäschke, and G. Stumme. Mining association rules in folksonomies. In V. Batagelj, H.-H. Bock, A. Ferligoj, and A. Žiberna, editors, *Data Science and Classification: Proc. of the 10th IFCS Conf.*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 261–270, Berlin, Heidelberg, 2006. Springer.

[30] P. Schmitz. Inducing ontology from Flickr tags. In *Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland*, May 2006.

[31] L. Zhang, X. Wu, and Y. Yu. Emergent semantics from folksonomies: A quantitative study. *Journal on Data Semantics VI*, 2006.

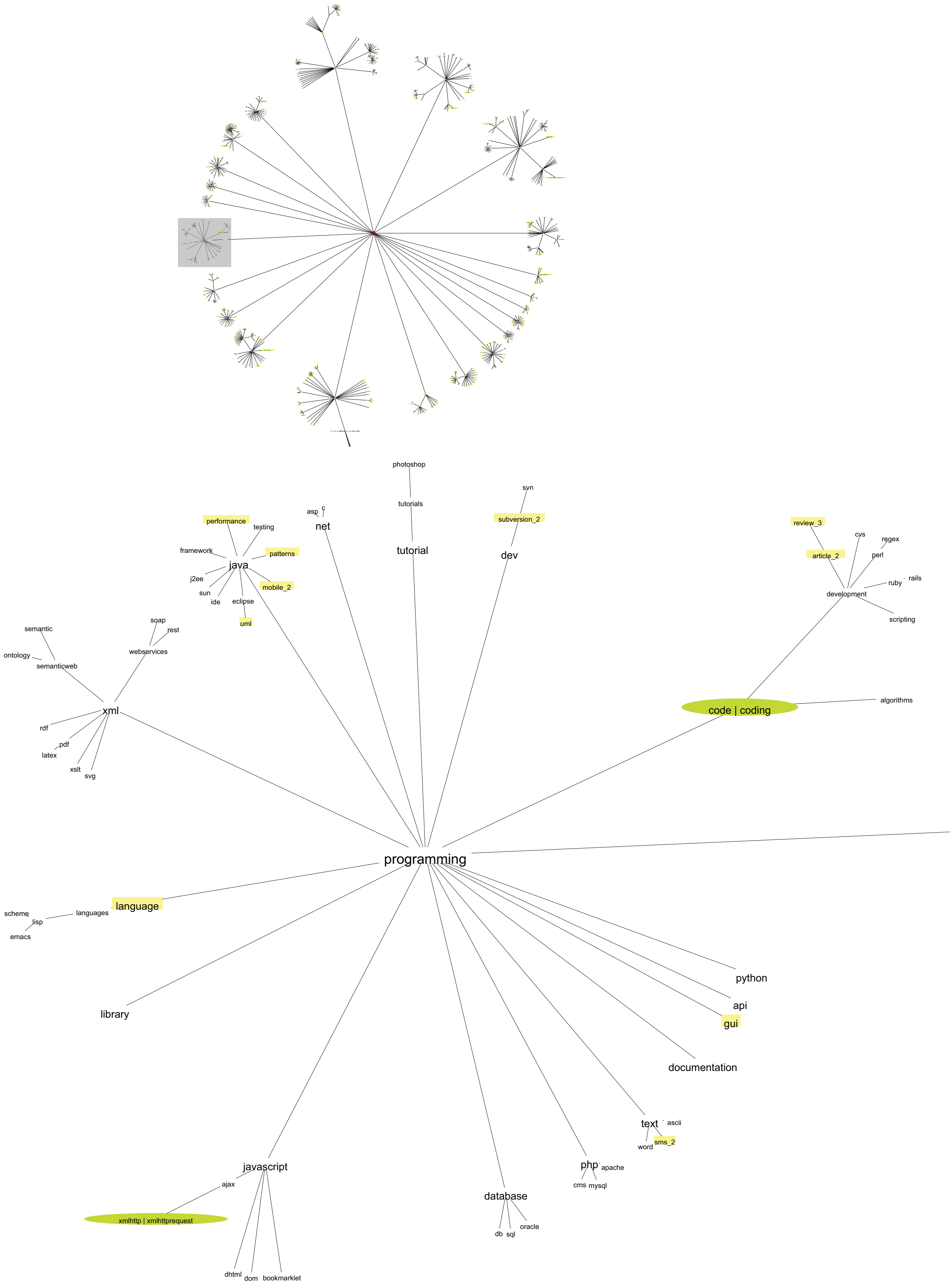


Figure 2: Excerpt (lower figure) from the learned ontology (upper figure) around the concept *programming*. Green nodes represent synsetized tags, while yellow nodes indicate ambiguous tags.

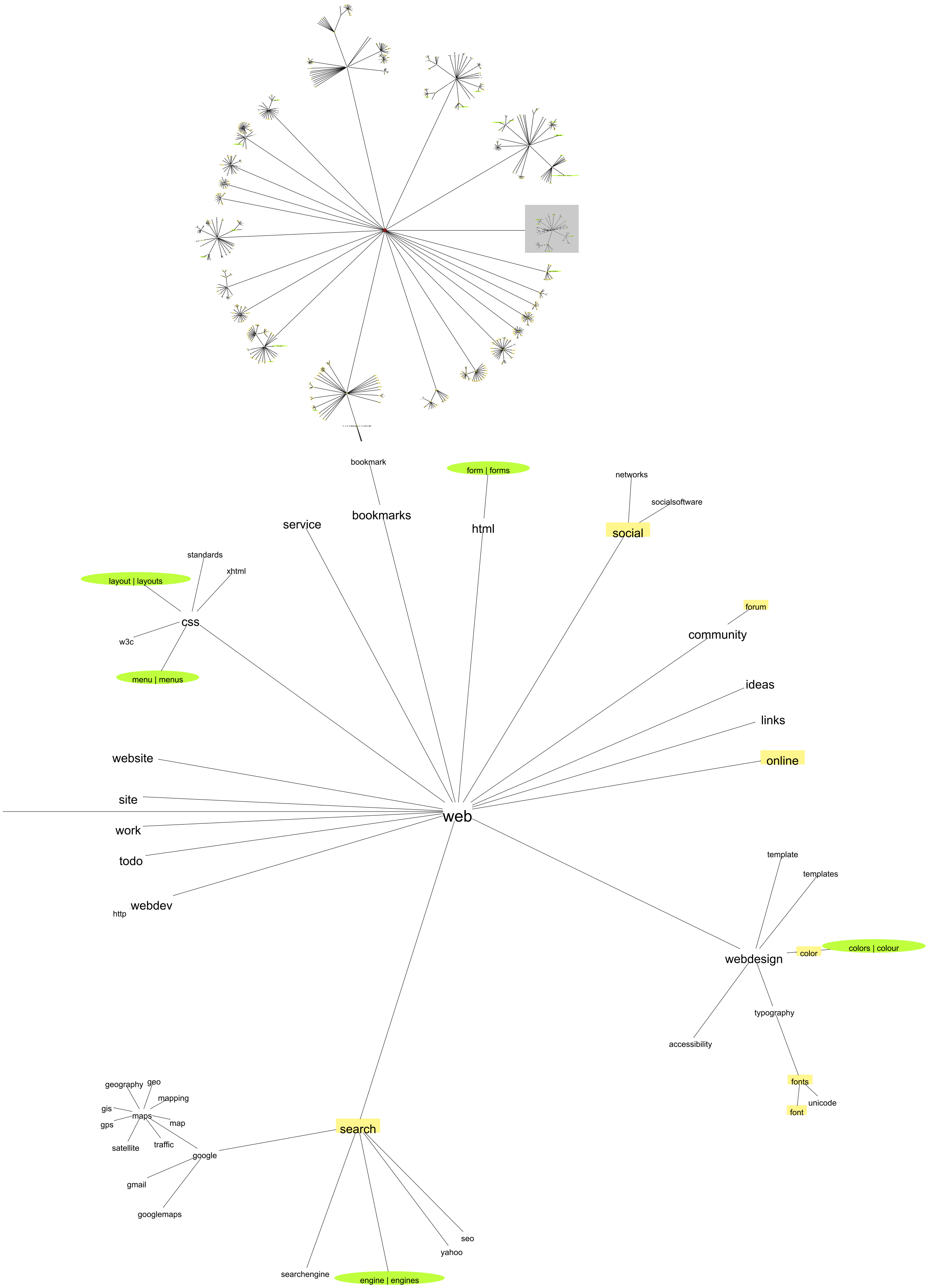


Figure 3: Excerpt (lower figure) from the learned ontology (upper figure) around the concept *web*. Green nodes represent synsetized tags, while yellow nodes indicate ambiguous tags.