The Horse before the Cart: Improving the Accuracy of Taxonomic Directions When Building Tag Hierarchies

Fahad Almoqhim, David E. Millard, Nigel Shadbolt

School of Electronics and Computer Science, University of Southampton, Southampton, UK

Abstract

Content on the Web is huge and constantly growing, and building taxonomies for such content can help with navigation and organisation, but building taxonomies manually is costly and time-consuming. An alternative is to allow users to construct folksonomies: collective social classifications. Yet, folksonomies are inconsistent and their use for searching and browsing is limited. Approaches have been suggested for acquiring implicit hierarchical structures from folksonomies, however, but these approaches suffer from the 'popularity-generality' problem, in that popularity is assumed to be a proxy for generality, i.e. high-level taxonomic terms will occur more often than low-level ones. To tackle this problem, we propose in this paper an improved approach. It is based on the Heymann–Benz algorithm, and works by checking the taxonomic directions against a corpus of text. Our results show that popularity works as a proxy for generality in at most 90.91% of cases, but this can be improved to 95.45% using our approach, which should translate to higher-quality tag hierarchy structures.

Keywords: Collective intelligence; folksonomies; tag hierarchies; social classification; social tagging.

1. Introduction

A common knowledge structure for organizing Web resources is taxonomy (Bloehdorn, 2008). However, as Web content today is huge and constantly growing, building and maintaining taxonomies for such content manually is costly and time consuming. Consequently, an alternative approach is to allow Web users to freely assign tags (descriptive metadata) to a Web content, and then produce a folksonomy (a set of user, tag, resource triples) as a result of that tagging (Vander Wal, 2007). However, folksonomies are beset by many problems – due to the lack of a consistent structure – such as synonyms, homonyms and ambiguity (Mathes, 2004; Guy and Tonkin, 2006; Benz, Hotho and Stutzer, 2010).

The lack of structure in folksonomies means that content retrieval tasks, such as searching, subscription and exploration, are limited (Begelman, Keller and Smadja, 2006; Limpens, Gandon and Buffa, 2008; Angeletou, Sabou and Motta, 2008; Lin and Davis, 2010). Acquiring latent hierarchical structures from folksonomies and creating tag hierarchies can be useful in improving content retrieval (Laniado, Eynard and Colombetti, 2007), and also other different tasks, such as building lightweight ontologies (Mika, 2007) and enriching knowledge bases (Zheng, Wu and Yu, 2008).

Although several promising approaches have been proposed as solutions to structure folksonomies, they come with limitations (Lin and Davis, 2010; Solskinnsbakk and Gulla, 2011). One of the most significant of these limitations is the popularity-generality problem. This arises from the tendency of tag-hierarchy building algorithms to use popularity as a proxy for generality. For example, if users tend to tag a picture of Riyadh attractions with ‘Riyadh’ much more than ‘KSA’, then ‘Riyadh’ will have a higher popularity and thus be located in a more general level than ‘KSA’ in spite of the fact that the relation makes more sense semantically if ‘KSA’ is the more general tag.

In previous work (Almoqhim, Millard and Shadbolt, 2014), we proposed an improved approach to building tag hierarchy to tackle the popularity-generality problem. Our approach combined and extended prior research in tag-hierarchy construction and lexico-syntactic patterns to correct the taxonomic direction between popular and more general tags. In this paper, we extend our previous work and show further improvement in building high-quality tag hierarchy.

2. Related Work

2.1. Learning concept hierarchy from text

Many works have been proposed for constructing concept hierarchies from unstructured text, for example Hearst (1992), Cimiano, Hotho and Staab (2005) and Snow, Jurafsky and Ng (2006). These works mainly exploit clustering techniques based on the distributional hypothesis (Harris, 1968), or methods based on lexico-syntactic patterns to acquire a certain semantic relationship to text, e.g. an 'is-a' or 'part-of' relationship (Cimiano, Hotho and Staab, 2005).

Although lexico-syntactic patterns can be useful to capture various semantic relations, the hyponym/hypernym relationship seems to yield the most accurate results, even with no pre-encoded knowledge. Moreover, they occur frequently in texts and across their genre boundaries (Hearst, 1992; Hearst, 1998). Table 1 shows Hearst's lexico-syntactic patterns (Hearst, 1992) that are commonly used for acquiring taxonomic relations from large text corpora (Cimiano, Hotho and Staab, 2005).

Table 1. Hearst's lexico-syntactic patterns for capturing hyponym/hypernym relations (Hearst, 1992).

No	*Pattern*	*Example*
1	**P** such as {C_1, C_2 ... , (and | or)} C_n	**European countries** such as *England* and *Spain*.
2	Such **P** as {C_1 ,} * {(or | and)} C_n	… works by such **authors** as *Herrick*, *Goldsmith*, and *Shakespeare*.
3	C_1 {, C_n} * {,} {(or | and)} other **P**	… *apple*, *orange*, *banana* or other **fruits**.
4	**P** {,} including {C_1,} * {or | and} C_n	… all **common-law countries**, including *Canada* and *England*.
5	**P** {,} especially {C_1,} * {or | and} C_n	… most **European countries**, especially *England*, *Spain*, and *France*.

2.2. Learning tag hierarchy from folksonomies

In recent years, several promising approaches have been proposed for constructing tag hierarchies from folksonomies. These approaches can be seen in three directions:

1. Clustering techniques-based approaches. First, tag pairs' similarities are calculated and then divided into groups based on these similarities. Then, the similarities of each two groups are calculated and then combined as one until all tags are in the same group.

2. Knowledge resources-based approaches. Several existing knowledge resources, such as Wikipedia and WordNet, can be used to discover the meaning of tags and their relationships.
3. Hybrid approaches. Some approaches of constructing tag hierarchies are based on the combination of both previous directions.

Table 3 summaries the learning tag hierarchy approaches reviewed in our work.

Table 2. Summary of the reviewed learning tag hierarchy approaches.

Approach	Class	Data Source	Brief description
(Heymann and Garcia-Molinay, 2006)	**Clustering techniques-based approaches**	Delicious and CiteULike	They use graph centrality in the tag-tag co-occurrence network to identify the generality order of the tags; i.e. the tag with the highest centrality is the most general tag thus it should be added to the tag hierarchy before others.
(Schmitz *et al.*, 2006)		Delicious	They use the theory of association rule mining to analyse and structure folksonomies.
(Schmitz, 2006)		Flickr	They adapt the work of Sanderson and Croft (1999) to introduce a subsumption-based model for building tag hierarchy.
(Mika, 2007)		Delicious	They present a graph-based model for constructing two tag hierarchies from folksonomies using statistical techniques. The first tag hierarchy is based on the overlapping set of user-tag networks, whereas the second is based on the overlapping set of object-tag networks.
(Benz, Hotho and Stutzer, 2010)		Delicious	They present an extension of the Heymann and Garcia-Molinay (2006) algorithm by applying tag co-occurrence as the similarity measure and the degree centrality as the generality measure. They succeed to produce clearer and more balanced tag hierarchies compared to the original algorithm.
(Plangprasopchok, Lerman and Getoor, 2010)		Flickr	They adapt affinity propagation proposed by Frey and Dueck (2007) to build deeper and denser tag hierarchies from folksonomies.
(Laniado, Eynard and Colombetti, 2007)	**Knowledge resources-based approaches**	Delicious	They use WordNet to disambiguate and structure the tags.
(Angeletou, Sabou and Motta, 2008)		Flickr	They present FLOR, an automatic approach for enriching folksonomies by linking them with related concepts in WordNet and online ontologies, using the Watson semantic search engine.
(Cantador *et al.*, 2008)		Delicious and Flickr	They introduce an approach that maps the tags with Wikipedia concepts, and then associates those tags with domain ontologies.
(Tesconi *et al.*, 2008)		Delicious	They use Wikipedia as an intermediate representation between the tags and some semantic resources, namely YAGO and WordNet.
(Garcia *et al.*, 2009)		Flickr	They propose an approach to disambiguate homonym tags through linking them to DBpedia entries.
(Specia and Motta, 2007)	**Hybrid approaches**	Delicious and Flickr	They present a semi-automatic approach relying on clustering techniques and using WordNet and Google to structure tags.
(Giannakidou *et al.*, 2008)		Flickr	They introduce a co-clustering approach for identifying the tag semantics by clustering tags, and relevant concepts from WordNet.
(Lin, Davis and Zhou, 2009)		CiteULike and Flickr	They propose an approach based on data mining techniques and WordNet concepts to discover the semantics in the tags.

2.3. Limitations of the current approaches

Using lexico-syntactic patterns for acquiring semantic relations from free text can provide reasonable precision, but their recall is low (Cimiano, 2006). Moreover, they are not suitable for acquiring

semantic relations in tag datasets as these datasets tend to be much more inconsistent than text corpus (Plangprasopchok, Lerman and Getoor, 2010).

Though several approaches based on clustering techniques have been proposed to structure folksonomies, they come with limitations (Lin and Davis, 2010; Solskinnsbakk and Gulla, 2011). One of these limitations is the suffering from the popularity-generality problem. On Flickr, for example, it was found (Plangprasopchok and Lerman, 2009) that the photos tagged with 'car' are ten times as that tagged with 'automobile'. And when applying clustering techniques, the tag 'car' is expected to have higher centrality, and then it will be perceived as more general than 'automobile'.

Knowledge resources, though, have been used to partially solve the limitations of clustering techniques approaches. Yet, such resources are limited and they can only cope with standard terms (Lin and Davis, 2010), whereas tag collections may contain spelling errors, idiosyncratic terms, abbreviations etc. Also, tags can be multi-lingual, which make these sources even harder to handle (Solskinnsbakk and Gulla, 2011).

In our work, we combine these approaches to take advantage of the accuracy of lexico-syntactic patterns, while keeping the scalability and flexibility of clustering techniques. We do this by using hyponym/hypernym patterns for correcting the direction of taxonomic tag pairs in a tag hierarchy generated by clustering techniques, hence addressing the popularity-generality problem.

3. Our Approach to Building High-Quality Tag Hierarchies

The detailed description of our original approach can be found in our previous work (Almoqhim, Millard and Shadbolt, 2014). The key aim of our approach is to solve the popularity-generality problem caused by using clustering techniques. To tackle this problem, our proposed approach extended a promising generality-based algorithm, based on Strohmaier *et al.* (2012), by using lexico-syntactic patterns applied to a large text corpus, i.e. English Wikipedia. The patterns that our approach used are a combination of the well-known Hearst's lexico-syntactic patterns (Table 1) and two other direct patterns: '*C* is a **P**' and '*C* is an **P**'. Figure 1 shows the process of our approach.

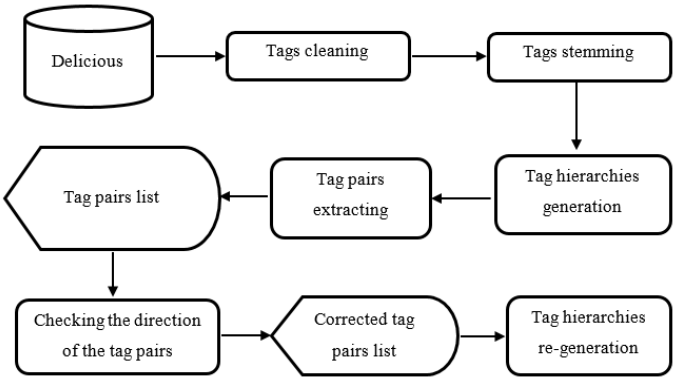

Figure 1. The process of our approach.

The algorithm we have developed and used in our approach (Table 3) is an extension of Benz's algorithm (Benz, Hotho and Stutzer, 2010), which itself is an extension of Heymann's algorithm (Heymann and Garcia-Molinay, 2006). Based on a comprehensive study of tag hierarchy construction algorithms, Strohmaier *et al.* (2012) show that generality-based approaches of tag hierarchy – with degree centrality as generality measure and co-occurrence as similarity measure, e.g. Benz's algorithm – show a superior performance compared to other approaches. Although we have used our original approach in this paper, the additional contribution here is a further improvement in building high-quality tag hierarchy. This improvement has been achieved by enhancing the evaluation methodology used in our previous work (this is explained further in Section 4.2).

Table 3. Pseudo-code for the proposed algorithm.

Input: *user-generated terms (tags)*
Output*: tag hierarchy*

1. Filter the tags by an occurrence threshold *occ*.
2. Order the tags in descending order by generality (measured by degree centrality in the tag–tag co-occurrence network).
3. Starting from the most general tag, as the root node, add all tags t*i* subsequently to an evolving tag hierarchy:
   (a) Calculate the similarities (using the co-occurrence weights as similarity measure) between the current tag t*i* and each tag currently present in the hierarchy, and append the current tag t*i* underneath its most similar tag *tag_sim*.
   (b) 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.
   (c) Check the taxonomic direction (t*i* → its suggested hypernym; i.e. *tag_sim* or the root) by using the proposed lexico-syntactic patterns, and calculate p_occ_1; i.e. in total, how many (t*i* → its suggested hypernym), with using the proposed patterns, found in Wikipedia.
   (d) Check the taxonomic direction (t*i* ← its suggested hypernym; i.e. *tag_sim* or the root) by using the proposed lexico-syntactic patterns, and calculate p_occ_2; i.e. in total, how many (t*i* ← its suggested hypernym), with using the proposed

patterns, found in Wikipedia.

(e) Correct the taxonomic direction if needed based on p_occ_1 and p_occ_2.

4. Apply a post-processing to the resulting hierarchy by re-inserting orphaned tags underneath the root node in order to create a balanced representation. The re-insertion is done based on Step 3.

4. Experimental Setup

To measure the performance of our methods, we applied both the Heymann–Benz approach and our own proposed approach, using five common tag similarity measures (Eq.1–Eq.5), to a large-scale folksonomy dataset collected from Delicious (Section 4.1), yielding 20 different tag hierarchies

$$Matching = |A \cap B| \quad \text{(Eq.1)}$$

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad \text{(Eq.2)}$$

$$Jaccard = \frac{|A \cap B|}{|A \cup B|} \quad \text{(Eq.3)}$$

$$Overlap = \frac{|A \cap B|}{\min(|A|, |B|)} \quad \text{(Eq.4)}$$

$$cosine = \frac{|A \cap B|}{\sqrt{|A| \times |B|}} \quad \text{(Eq.5)}$$

where ‘A’ is the folksonomies set that contains *Tag a*, and ‘B’ is the folksonomies set that contains the co-occurrence *Tag b*.

4.1. Datasets

In our experiments, we have used two large datasets:

1. Delicious dataset. This is a large-scale folksonomy dataset from the PINTS experimental dataset (PINTS, 2006–2007) containing a systematic crawl of Delicious during 2006 and 2007. Table 4 summarised the statistics of the dataset.

Table 4. Statistics of the Delicious dataset.

Dataset	*Users*	*Tags*	*Resources*	*Tag assignments*
Delicious	532,924	2,481,698	17,262,480	140,126,586

2. Wikipedia dataset. To tackle the popularity-generality problem by using the proposed lexico-syntactic patterns, we have chosen Wikipedia dataset since it is currently the largest knowledge

repository available on the Web. The dataset that we have extracted contains 4,487,682 English Wikipedia articles (as collected in March 2014).

4.2. Evaluation methodology

To test the proposed approach to building tag hierarchy against the Heymann–Benz approach, we have selected WordNet (Miller, 1995) dataset for three reasons:

1. It is considered to be a gold-standard dataset for evaluating hyponym/hypernym relations (Snow, Jurafsky and Ng., 2004).
2. WordNet is a reasonable reference for our purpose, i.e. solving the popularity-generality problem, as a significant fraction of the popular tags in Delicious is covered by WordNet (Almoqhim, Millard and Shadbolt, 2014).
3. We needed to avoid any reference dataset that was constructed automatically or based on Wikipedia as we have used it in our approach.

WordNet is a structured lexical database of the English language that is created manually by experts. It contains 206,941 terms grouped into 117,659 synsets (WordNet, 2014). The synsets are connected by several lexical relations, but the most frequent of these relations is the hyponym/hypernym relation.

Hyponym/hypernym relations are transitive; e.g. if ‘armchair’ is a kind of ‘chair’, and ‘chair’ is a kind of ‘furniture’, then ‘armchair’ is a kind of ‘furniture’. By extracting all transitive hyponym/hypernym relations in WordNet we were able to evaluate our approach with a more reasonable size of taxonomy reference, as shown in Section 5, than we did before. Whereas in previous work (Almoqhim, Millard and Shadbolt, 2014) we had extracted 364,135 direct taxonomic terms among synsets in WordNet, we have, for this paper, extracted 2,153,520 direct and inherited taxonomic terms among synsets in WordNet.

5. Results and Analysis

We have run our experiment twice. First, we applied the Heymann–Benz algorithm and our proposed algorithm, using all the five mentioned tag similarity measures, to the Delicious dataset, to build ten tag hierarchies. Second, we reran the experiment but without using a tag similarity threshold that was proposed by the Heymann–Benz algorithm. Then, we evaluated the direction accuracy of all the produced taxonomic tag pairs from the two experiments against WordNet. Table 5 illustrates a few examples of these produced taxonomic tag pairs, using the five selected similarity measures, whereas Table 6 shows the WordNet coverage of these taxonomic tag pairs.

Table 5. Examples of produced tag pairs for each of the selected similarity measures.

Measure	*Rank*	*Tag A*	*Tag B*	*Rank*	*Tag A*	*Tag B*
Matching	1	**Blog**	design	1000	**Display**	technology
Dice			design			Lcd
Jaccard			design			Lcd
Overlap			blogger beast			Tft
Cosine			Daily			lcd
Matching	100	**Daily**	Blog	5000	**Maple**	php
Dice			News			willow
Jaccard			News			willow
Overlap			Blog			willow
Cosine			News			willow
Matching	500	**Weather**	News	10000	**Bridesmaid**	dress
Dice			forecast			bridal
Jaccard			forecast			bridal
Overlap			Noaa			dress
Cosine			Forecast			bridal

Table 6. WordNet coverage of tags in produced hierarchies.

	Matching	*Dice*	*Jaccard*	*Overlap*	*Cosine*
WordNet coverage	48.37%	50.77%	50.77%	49.37%	50.70%

Having established this, the next step was to compare the tag pair directions produced by the Heymann–Benz algorithm and our algorithm against the taxonomic directions listed in WordNet. This was to allow us to measure how many times generality was a successful proxy for popularity in the Heymann–Benz algorithm, and also the extent to which our approach improves on this.

Table 7 shows the results. The first observation that can be drawn from the results is that while popularity-generality has been identified as a flaw of clustering approaches, using this assumption the Heymann–Benz algorithm is moderately successful, with at most 90.91% in all cases.

Table 7. Taxonomic tag pairs evaluation, using selected similarity measures and a similarity threshold for each measure, against WordNet.

	% Agreement with WordNet	
	Heymann–Benz algorithm	**Our algorithm**
Matching	91.39%	94.62%
Dice	33.33%	83.33%
Jaccard	36.36%	81.81%
Overlap	**90.47%**	**95.23%**
Cosine	62.50%	87.50%

The second observation that can be drawn from Table 7 is that there is a considerable improvement achieved by our proposed algorithm compared to the Heymann–Benz algorithm amongst all the mentioned tag similarity measures. This means, regardless of the similarity measure, our approach has succeeded in correcting the direction of taxonomic tag pairs that were generated in the wrong direction by the Heymann–Benz algorithm. In the best case (Overlap), this leads to a correct direction in more than 95% of cases. Another observation from the results is that, amongst all the tag similarity measures under study, the Overlap measure achieves the best performance of producing taxonomic tag pairs against WordNet, while the Matching measure achieves the biggest amount of produced tag pairs found in WordNet regardless of the taxonomic direction.

Table 8. Taxonomic tag pairs evaluation, using selected similarity measures and without using a similarity threshold, against WordNet.

	% Agreement with WordNet	
	Heymann–Benz algorithm	**Our algorithm**
Matching	90.82%	94.90%
Dice	40.00%	93.33%
Jaccard	27.27%	90.91%
Overlap	**90.91%**	**95.45%**
Cosine	65.38%	88.46%

Table 8, however, shows the results of rerunning the experiment without using a tag similarity threshold, but with all selected similarity measures. Besides the previous observations in Table 7, Table 8 validates that without using a similarity threshold, as proposed by the Heymann–Benz algorithm, both the Heymann–Benz algorithm and ours are able to generate more taxonomic tag pairs that can be found in WordNet.

6. Conclusion

Folksonomies have recently emerged as an alternative approach to traditional classifications of organizing Web content. Yet, their lack of consistent structure leads to many problems, such as synonyms, homonyms and ambiguity. Hence many approaches have been offered to solve these issues by proposing methods for acquiring latent hierarchical structures from folksonomies and building tag hierarchies. However, these approaches come with limitations, one of the most important of which is the popularity-generality problem, where it (sometimes inaccurately) assumes that since a tag occurs more frequently it must be more general and hence appear higher in the hierarchy. Thus, we have presented an experiment to test this assumption, and introduced an approach to reduce its impact.

Our proposed approach extends a promising generality-based algorithm by using lexico-syntactic patterns for capturing hyponym/hypernym relations with the purpose of distinguishing between popular and general tags. For this purpose, we have used Wikipedia as the text corpus, and for evaluation we have used WordNet as a gold-standard reference.

Our results show that popularity acts as a successful proxy for generality in 27% to 90% of cases, depending on the similarity measure used, and that, among all the selected tag similarity measures, the performance of our proposed algorithm outperforms the Heymann–Benz algorithm (correct in between 81% and 95% of cases). This means, regardless of the similarity measure, our approach has succeeded in correcting the direction of taxonomic tag pairs that were wrongly created by the Heymann–Benz algorithm. This improvement will result in building higher-quality tag hierarchy structure and semantics. Also found was that in terms of the selected tag similarity measures, the Overlap measure achieves the best performance of producing taxonomic tag pairs against WordNet. Finally, we have revealed that ignoring the similarity threshold, in both the Heymann–Benz algorithm and ours, results in better taxonomic tag pairs, in terms of quantity and quality.

For future work, we plan to investigate which lexico-syntactic patterns are most successful in correcting errors, and whether any introduce significant errors. Also, based on the results we have shown, we are going to use a dynamic knowledge repository, such as a search engine, rather than a

static knowledge resource, like Wikipedia. This should help in increasing the coverage and occurrences of the tags in any tag dataset.

References

Almoqhim, F., Millard, D.E. and Shadbolt, N. (2014). Improving on Popularity as a Proxy for Generality When Building Tag Hierarchies from Folksonomies, *6th International Conference on Social Informatics (SocInfo 2014)*, 95–111.

Angeletou, S., Sabou, M. and Motta, E. (2008). Semantically Enriching Folksonomies with FLOR, *1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008)*, Tenerife, Spain.

Begelman, G., Keller, P. and Smadja, F. (2006). Automated Tag Clustering: Improving Search and Exploration in the Tag Space, *Collaborative Web Tagging Workshop at WWW2006*, 15–33.

Benz, D., Hotho, A. and Stutzer, S. (2010). Semantics Made by You and Me: Self-Emerging Ontologies Can Capture the Diversity of Shared Knowledge, *2nd Web Science Conference (WebSci10)*, Raleigh, North Carolina, USA.

Bloehdorn, S. (2008). Kernel Methods for Knowledge Structures, Ph.D. diss., Karlsruhe Institute of Technology, Karlsruhe, Germany.

Cantador, I., Szomszor, M., Alani, H., Fernández, M. and Castells, P. (2008). Enriching Ontological User Profiles with Tagging History for Multi-Domain Recommendations, *1st International Workshop on Collective Semantics: Collective Intelligence and the Semantic Web (CISWeb 2008)*, Tenerife, Spain.

Cimiano, P. (2006). *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*, Springer, New York, USA.

Cimiano, P., Hotho, A. and Staab, S. (2005). Learning Concept Hierarchies from Text Corpora Using Formal Concept Analysis, *Journal of Artificial Intelligence Research*, **24**, 305–339.

Frey, B.J. and Dueck, D. (2007). Clustering by Passing Messages between Data Points, *Science*, **315**, 972–976.

Garcia, A., Szomszor, M., Alani, H. and Corcho, O. (2009). Preliminary Results in Tag Disambiguation Using DBpedia, *1st International Workshop in Collective Knowledge Capturing and Representation*, California, USA.

Giannakidou, E., Koutsonikola, V., Vakali, A. and Kompatsiaris, Y. (2008). Co-Clustering Tags and Social Data Sources, *9th International Conference On Web-Age Information Management*, 317–324.

Guy, M. and Tonkin, E. (2006). Tidying up Tags, *D-Lib Magazine*, **12**.

Harris, Z.S. (1968). *Mathematical Structures of Language*, Wiley and Sons, New York, USA.

Hearst, M.A. (1992). Automatic Acquisition of Hyponyms from Large Text Corpora, *14th conference on Computational Linguistics*, 539–545.

Hearst, M.A. (1998). 'Automated discovery of wordnet relations,' in Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*, MIT Press, Cambridge, USA.

Heymann, P. and Garcia-Molinay, H. (2006). *Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems*, Stanford InfoLab, Stanford, California, USA.

Laniado, D., Eynard, D. and Colombetti, M. (2007). Using WordNet to Turn a Folksonomy into a Hierarchy of Concepts, *4th Italian Semantic Web Workshop: Semantic Web Application and Perspectives*, 192–201.

Limpens, F., Gandon, F. and Buffa, M. (2008). Bridging Ontologies and Folksonomies to Leverage Knowledge Sharing on the Social Web: a Brief Survey, *1st Workshop on Social Software Engineering and Applications (SoSEA)*, 13–18.

Lin, H. and Davis, J. (2010). Computational and Crowdsourcing Methods for Extracting Ontological Structure from Folksonomy, *7th Extended Semantic Web Conference (ESWC'10)*, 472–477.

Lin, H., Davis, J. and Zhou, Y. (2009). An Integrated Approach to Extracting Ontological Structures from Folksonomies, *6th European Semantic Web Conference*, 654–668.

Mathes, A. (2004). Folksonomies-Cooperative Classification and Communication through Shared Metadata, *Computer Mediated Communication*, **47**.

Mika, P. (2007). Ontologies are Us: a Unified Model of Social Networks and Semantics, *Web Semantics: Science, Services and Agents on the World Wide Web*, **5**, 5–15.

Miller, G.A. (1995). WordNet: a Lexical Database for English, *Communications of the ACM*, **38**, 39–41.

PINTS. 2006–2008. *PINTS – Experimental Data Sets*. Available at: http://www.uni-koblenz-landau.de/koblenz/fb4/AGStaab/Research/DataSets/PINTSExperimentsDataSets/index.html [Accessed 10 April 2014].

Plangprasopchok, A. and Lerman, K. (2009). Constructing Folksonomies from User-Specified Relations on Flickr, *18th International World Wide Web Conference*, 781–790.

Plangprasopchok, A., Lerman, K. and Getoor, L. (2010). Growing a Tree in the Forest: Constructing Folksonomies by Integrating Structured Metadata, *16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 949–958.

Sanderson, M. and Croft, B. (1999). Deriving Concept Hierarchies from Text, *22nd ACM Conference of the Special Interest Group in Information Retrieval*, 206–213.

Schmitz, P. (2006). Inducing Ontology from Flickr Tags, *Collaborative Web Tagging Workshop at WWW2006*, Edinburgh, Scotland.

Schmitz, C., Hotho, A., Jäschke, R. and Stumme, G. (2006). Mining Association Rules in Folksonomies, *10th IFCS Conference: Studies in Classification, Data Analysis and Knowledge Organization*, 261–270.

Snow, R., Jurafsky, D. and Ng, A.Y. (2006). Semantic TaxonomyIinduction from Heterogenous Evidence, *21st International Conference on Computational Linguistics*, 801–808.

Snow, R., Jurafsky, D. and Ng, A.Y. (2004). Learning Syntactic Patterns for Automatic Hypernym Discovery, *18th Annual Conference on Neural Information Processing Systems (NIPS 2004)*, Vancouver, Canada.

Solskinnsbakk, G. and Gulla, J.A. (2011). Mining Tag Similarity in Folksonomies, *3rd International Workshop on Search and Mining User-Generated Contents (SMUC '11)*, 53–60.

Specia, L. and Motta, E. (2007). Integrating Folksonomies with the Semantic Web, *4th European Conference on the Semantic Web: Research and Applications*, 624–639.

Strohmaier, M., Helic, D., Benz, D., Körner, C. and Kern, R. (2012). Evaluation of Folksonomy Induction Algorithms, *ACM Transactions on Intelligent Systems and Technology*, **3**.

Tesconi, M., Ronzano, F., Marchetti, A. and Minutoli, S. (2008). Semantify del.icio.us: Automatically Turn your Tags into Senses, *Social Data on the Web Workshop at the 7th International Semantic Web Conference*, Karlsruhe, Germany.

Vander Wal, T. 2007. *Folksonomy Coinage and Definition*. Available at: http://vanderwal.net/folksonomy.html [Accessed 10 April 2014].

WordNet. 2014. *WordNet 3.0 Database Statistics*. Available at: http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html [Accessed June 2014].

Zheng, H., Wu, X. and Yu, Y. (2008). Enriching WordNet with Folksonomies, *12th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD'08)*, 1075–1080.