LAMP-TR-155
CS-TR-4980
UMIACS-TR-2011-07

May 2011

TAMING SOCIAL TAGS: COMPUTATIONAL LINGUISTIC ANALYSIS OF TAGS FOR IMAGES IN MUSEUMS

Judith L. Klavans[1], Raul Guerra[1], Rebecca LaPlante[1], Robert Stein[2], and Edward Bachta[2]

[1]Computational Linguistics and Information Processing
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742-3275
{jklavans, laplante}@umd.edu, raulguerrap@gmail.com

[2]Indianapolis Museum of Art
4000 Michigan Road
Indianapolis, IN 46208-3326
{rstein, ebachta}@imamuseum.org

Abstract

This paper reports on the linguistic analysis of a tag set of nearly 50,000 tags collected as part of the steve.museum project. The tags describe images of objects in museum collections. We present our results on morphological, part of speech and semantic analysis. We demonstrate that deeper tag processing provides valuable information for organizing and categorizing social tags. This promises to improve access to museum objects by leveraging the characteristics of tags and the relationships between them rather than treating them as individual items. At a high level, the paper shows the value of using computational linguistic techniques in interdisciplinary projects with museums and libraries.

Keywords: social tag, morphological analysis, part of speech analysis, semantic analysis, image description

The support of this research by the Institute of Museum and Library Services under grant LG-30-08-0117-08 is gratefully acknowledged.

Taming Social Tags: Computational Linguistic Analysis of Tags for Images in Museums

Judith L. Klavans[1], Raul Guerra[1], Rebecca LaPlante[1], Robert Stein[2], and Edward Bachta[2]
[1]University of Maryland College Park, [2]Indianapolis Museum of Art
{jklavans, laplante}@umd.edu, raulguerrap@gmail.com,
{rstein, ebachta}@imamuseum.org

Abstract

This paper reports on the linguistic analysis of a tag set of nearly 50,000 tags collected as part of the steve.museum project. The tags describe images of objects in museum collections. We present our results on morphological, part of speech and semantic analysis. We demonstrate that deeper tag processing provides valuable information for organizing and categorizing social tags. This promises to improve access to museum objects by leveraging the characteristics of tags and the relationships between them rather than treating them as individual items. At a high level, the paper shows the value of using computational linguistic techniques in interdisciplinary projects with museums and libraries.

1 Challenges of Tags

We address two of the questions of this workshop: first, we consider which types of linguistic analysis are useful in discriminating the noise from social tags, for example from multiple user input, unconstrained language, etc.; and second, we consider the issue of semantic categorization of tags for understanding this social data using domain models, for example medical, legal, or (in this case) the art history domain.

Identifying linguistic traits of tags provides some unique challenges. Linguistic analysis of words or phrases within the context of full text draws upon context to provide clues to the characteristics of the individual components. With tags, especially those affiliated with images, the contextual environment is minimal or non-existent.

The fundamental research questions driving this research are:

(1) How can a set of social tags describing images be analyzed using (a) computational linguistic tools, such as morphological analyzers, part of speech taggers; (b) online lexical resources such as WordNet (Miller, 1995) or the Art and Architecture Thesaurus (Getty, 2010), and (c) clustering (Becker et al., 2010) to characterize an image?

(2) What are the optimal linguistic processes to normalize tags since these steps could have a large effect on later processing (e.g., clustering)?

(3) In what ways can social tags be associated with other information to improve users' access to museum objects?

Tags are but one of the many information sources for museum object metadata. The environment for analyzing words and phrases as part of running text is the entire body of text. The analogous environment for tags associated with museum objects or with images of those objects is the tag or tag phrase itself and the set of tags associated with a given image. We refer to this as the tag's "internal environment."

Whereas an individual social tag's "internal environment" refers to all other tags for that object, there is a larger set of surroundings in which to view each tag. We refer to this as the social tag's "external environment", distinct from the "internal environment." This larger milieu consists of related information such as a gallery label (which typically gives basic information on the name of the work, the creator, the creation date, the medium, and ownership), text written about that image in a museum catalog or publication, or even (for collections such as Flickr) image descriptions such as "Susie's birthday party at the Eiffel Tower."

Furthermore, social tags may reflect information above and beyond what is visible in the image, such as "painted in the style of Goya", or "taken by Susie's sister". Lastly, the type of objects (painting, furniture, ceramic) for which the tag was assigned may provide information to identify its linguistic traits. Similarly, the word "wind" when used in the context of a collection of photograph of birds in flight (say in an Audobon collection) would have a stronger likelihood of being the noun form than if it was used in a collection of images related to watch-making.

2 Related Work

Social tags have been computationally analyzed from many perspectives. One of the most useful of these perspectives has been for applications such as product reviews (Pang and Lee 2008). However, since the focus of this paper is on tagging of objects and images, we concentrate on this subset of related research.

Perhaps the most well known image-tagging application for the larger community is Peekaboom, presented in von Ahn et al. (2006). The purpose of this project is to gather user-generated input for locating objects in images to train computer vision algorithms. The number and types of tags collected is rich and varied. However, unlike our project, the images are harvested from Web pages and contain little associated metadata, as images of an object from a library or museum would. In addition, Peekaboom focuses on a literal interpretation of a tag rather than exploring more abstract concepts or personal interpretations for enhanced image access. Similarly, the Visual Dictionary Project (VDP) (Torralba et al., 2008) has collected user input to create training data for vision recognition systems (http://groups.csail.mit.edu/vision/TinyImages/).
Like the VDP, data collected from the steve.museum project is available for community use. The combination of visual features and tags (Aurnhammer et al., 2006) is related in that tags need to be analyzed in terms of their semantics. Begelman et al. (2006) explore the use of clustering over tags for the same image to identify semantically related tags and thus help users in the tagging experience. This research is relevant to the tag collection process

Unlike other image tagging projects, such as Flickr or Google Image Labeler or Tag Cow, the data in this project was collected within a highly controlled environment over a carefully selected set of images with participation from 18 museum partners interested in the use of social media for museums, a rapidly growing area of interest for the museum community (Vogel, 2011).

Stewart (2010) explores the variety of indexing approaches for subject descriptions of images of historical importance, emphasizing that social tagging offers alternatives to institutionally mediated access and a shift in the locus of control that highlights the different or competing interpretations of image content that are available. The study demonstrates the wide variety of tags expected even over a representational image, but the linguistic properties of tags is not addressed. Lee and Schleyer (2010) examine the mapping between tags and the controlled subject headings from MeSH terms, showing (as did Trant, 2007) that there is little overlap. Lee and Schleyer (op. cit) use the Porter Stemmer for normalization (van Rijsbergen et al., 1980) and Google tools for spell-checking and compound word separation, which was adequate for their preprocessing needs. In our research, we are taking this analysis a step further to examine parts of speech, semantic categorization and disambiguation.

3 Computational Linguistic Tools and Tag Analysis

3.1 Description of the Tag Data:

The steve project (http://www.steve.museum) is a multi-institutional collaboration exploring applications of tagging in museums. The project seeks to improve collection documentation and public access to online object collections. To do so, project partners engage the general public to contribute tags describing objects in their own words. Initial research showed that user tags in the steve.museum project enhance existing object documentation, providing information that is not currently recorded in the museums' formal documentation (Trant and Wyman, 2006; Trant et al., 2007; Trant et al., 2009). The T3: Text, Tags, Trust project (http://t3.umiacs.umd.edu) is building on this research and developing open source software that applies techniques from computational linguistics that enhance the usability of the collected set of social tags.

Another important contribution of the initial steve.museum research project is an original dataset of nearly 50,000 tags applied to 1,785 works. This dataset is the basis of the research described in this paper. Trant et al. (2009) describes the analysis of the tags collected by token. Further analysis by Klavans et al. (2011) extended this analysis to include an examination of one, two, and three word tags by type and token. The resulting analysis, described below, reveals significantly different results.

3.2 Morphological Analysis

Klavans et al. (2011) explore various processes needed to normalize the steve.museum dataset in a pipeline architecture. These preprocessing techniques include handling the range of anomalous characters occurring in tags, among them white spaces, character returns, and punctuation. Analysis showed that it may be desirable to conflate tags related to the same topic rather than counting them as distinct tags. The Morphy lemmatizer, an element of the Natural Language Toolkit (NLTK) (Bird et al., 2009), was used to conflate tags from the dataset. Other stemmers and lemmatizers were tested but Morphy provided the best results as reported in earlier work. In addition, NLTK is a Creative Commons-licensed open source project.

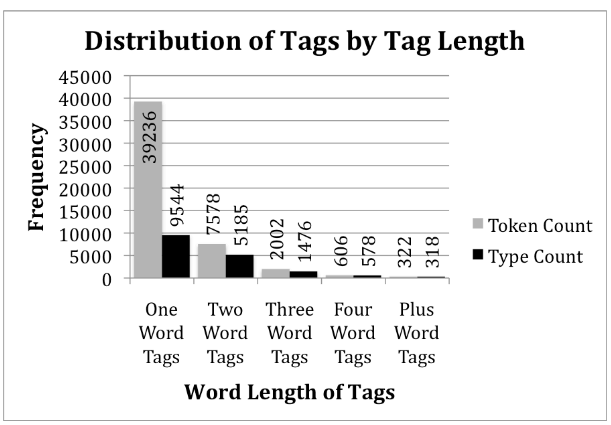

Figure 1. Distribution of tags by tag length.

Figure 1 demonstrates that simple preprocessing to conflate tags dramatically reduces the number of tags by type compared with the number of tags by token.

As shown in Figure 1, the majority of tags (79% by token, 52% by type) consist of one word, followed by those consisting of two words (15% by token, 33% by type). Only a small percentage of tags (6% by token, 15% by type) are longer than 2 words. Since basic tag frequency is used by many subsequent analyses, the impact of conflation cannot be underestimated.

3.3 Part of Speech Analysis

One of the original contributions of this paper is to provide insight on the role of part of speech (POS) tagging in tag normalization and analysis. Operations like morphological analysis may in fact depend on POS tagging as discussed below. Similarly, domain-dependent factors may influence standard POS tagging approaches.

Many words are ambiguous in their part of speech; part of speech tagging has long been a part of the computational linguistic analysis process in order to perform higher level analysis, such as parsing and phrase identification. Most POS taggers use lexical probability, trained on a particular corpus, along with contextual probability derived from context. We used the NLTK's bigram tagger that takes as input the current word together with the POS of the previous word. However, because most tags are single words, as shown in Figure 1, we have had to default the bigram tagger to a unigram tagger. This unigram tagger also then defaults to a tagger that tags everything as a noun (NN), since this is the most frequently occurring POS tag. All taggers were trained with the Brown Corpus.

The frequency of tags was calculated by performing a very simple normalization. Symbols were removed from the left and right of a word. Tags were then tokenized. Stop words and symbols were removed from tags. If the tag was an empty string because of removing symbols, then the original tag was removed (since it was rendered an empty string.) Afterward, the POS tagger was used to obtain the best-guess POS tag for each word in the tag. These POS tags are then used in conjunction with the NLTK's lemmatizer to output the base form of each word in the tag.

It is important to note that different preprocessing and normalization methods yield slightly different output. In this paper, in contrast to Klavans (2011), cardinal and ordinal numbers are retained in the tag data set; typically, numbers are removed from social tag sets, but in this case, they reflect important information about an object. For example, "17thC", "1400 B. C." or "3 sisters." We have used two approaches to POS tagging for comparison. First, we used NLTK's pretrained classifier-based tagger using a maximum entropy classifier and trained on the Penn Treebank corpus which consists of the Brown Corpus, 1 million words of the Wall Street Journal (WSJ), and additional data from spoken language (http://www.cis.upenn.edu/~treebank/). This means that we utilized the predefined features that came with this classifier. The maximum entropy classifier in NLTK extracts predefined features and utilizes a linear combination of them to classify the observed word into a POS category. Because it is a probabilistic classifier it also assigns a probability for each POS tag and chooses the one with the highest probability. To evaluate, we extracted 20 instances of the most frequent categories, in this case NN, JJ, NNS, NN-NN, JJ-NN, and VBG and manually examined them for accuracy. We observed that approximately 84% of the tags were correct.

We then tried the Brown trained POS tagger, which performed better. This was somewhat surprising since the Brown corpus is a subset of the Penn Treebank. Our manual examination of results shows over 95% correct POS assignments. We attribute this difference to the fact that social tags are generally taken from a non-esoteric vocabulary; the addition of WSJ data for the Maximum Entropy classification may in fact be higher performing for edited text, whereas the Brown-trained POS tagger seems to be more tuned to the type of language social tags reflect. We have created an randomly selected gold standard of 850 items from the 50,000 steve tagset, marked for POS using the Penn TB tags, but since the Brown tags reflect a different tagset, we will need to adopt our methodology to obtain a reliable comparison.

The results of these different attempts have brought us closer to the answer of one of the fundamental research questions driving this project: to figure out how to best handle the normalization of tags since this could impact basic statistical issues, such as frequency values. Further down the analysis pipeline, processes such as clustering and similarity detection rely on frequency.

Note that one of the major challenges of POS tagging of the dataset is that most items are one word (e.g. "blue", "wind", "squares".) As a result, there is little information in a tag itself to help decipher the nature of the words within that short string. Other tags on the same object may provide some context. For example, "blue" in the context of "sad" or "lonely" indicate the meaning of "blue" as "saddened"; the example of part of speech for "wind" was given above.

However, since tags can reflect a wide variety of characteristics, such as subject matter (woman), biographical data (painted by Pablo Picasso), or opinion (scary), there may be a loose relationship between an individual tag and the set of tags on the same object. For example, "sad" and "lonely" might apply to one of Picasso's blue period paintings, which are predominantly blue in color. There is no unambiguous way of knowing which sense of "blue" is intended.

Once all the tags have been assigned a POS, then an analysis of patterns can be performed. Figure 2 shows a graph with the frequency of each POS pattern in the tag set based on token count. The x-axis reflects the order of the POS patterns by frequency of occurrence where 1 (NN) is the most frequently occurring pattern and the least frequently unique patterns are from 409th place to 1253rd. The y-axis reflects the frequency of each pattern. Not surprisingly, the most frequent tag is NN, singular noun, for 25205 of the single word tags, followed by JJ, adjective (n=6319) and the NNS, plural noun (n=4041). The next most frequent patterns are for two word phrases, NN-NN, noun-noun compound, and then JJ-NN, adjective-noun. Again, given the context of museum objects and images of these objects, this is to be expected. At the same time, a deeper analysis of results is needed to confirm that labeling is as expected, since typically noun compounds in English are ambiguous. The next category is VBG, which are gerunds such as "sitting" or "beating". Our initial examination of these VBG's shows that approximately 60% are used as nominals, but this is the focus of future research. Similarly, VBN's are usually used adjectivally, so that the nominal VBG's could be conflated with NN's and VBN's with JJ's. Proper nouns, ordinal numbers with nouns, and

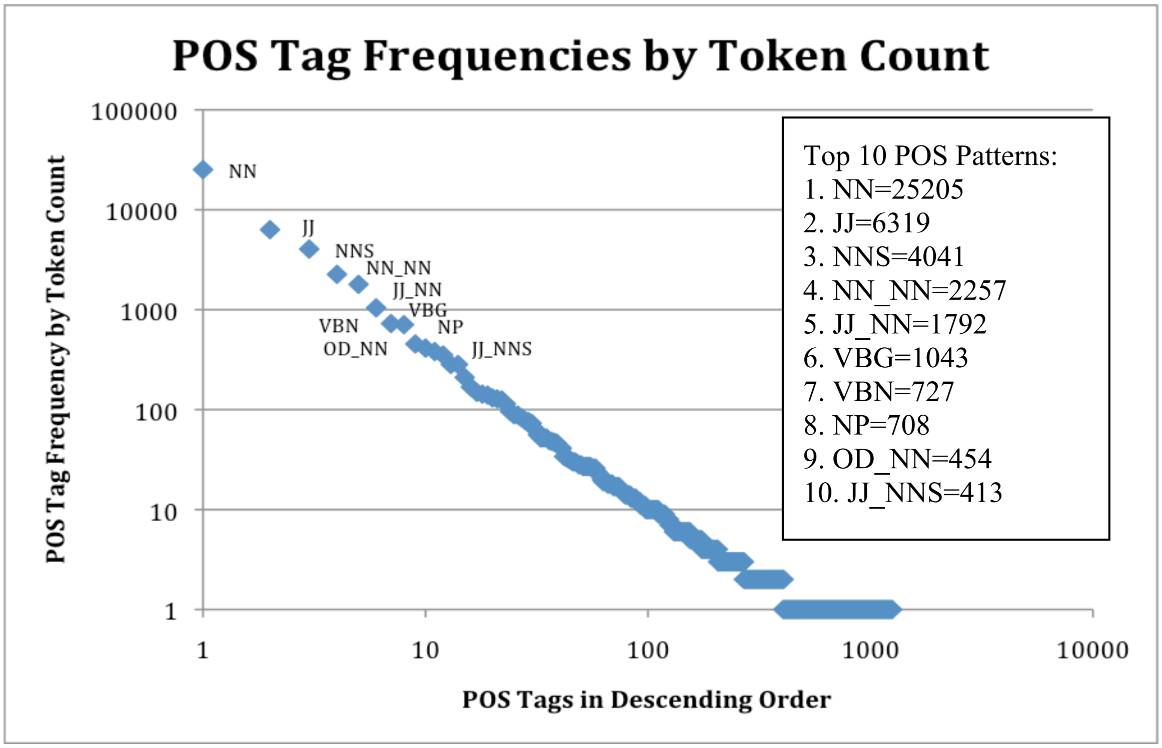

Figure 2. Part of Speech (POS) tag frequencies by token count.
Note that each axis is on a logarithmic scale.

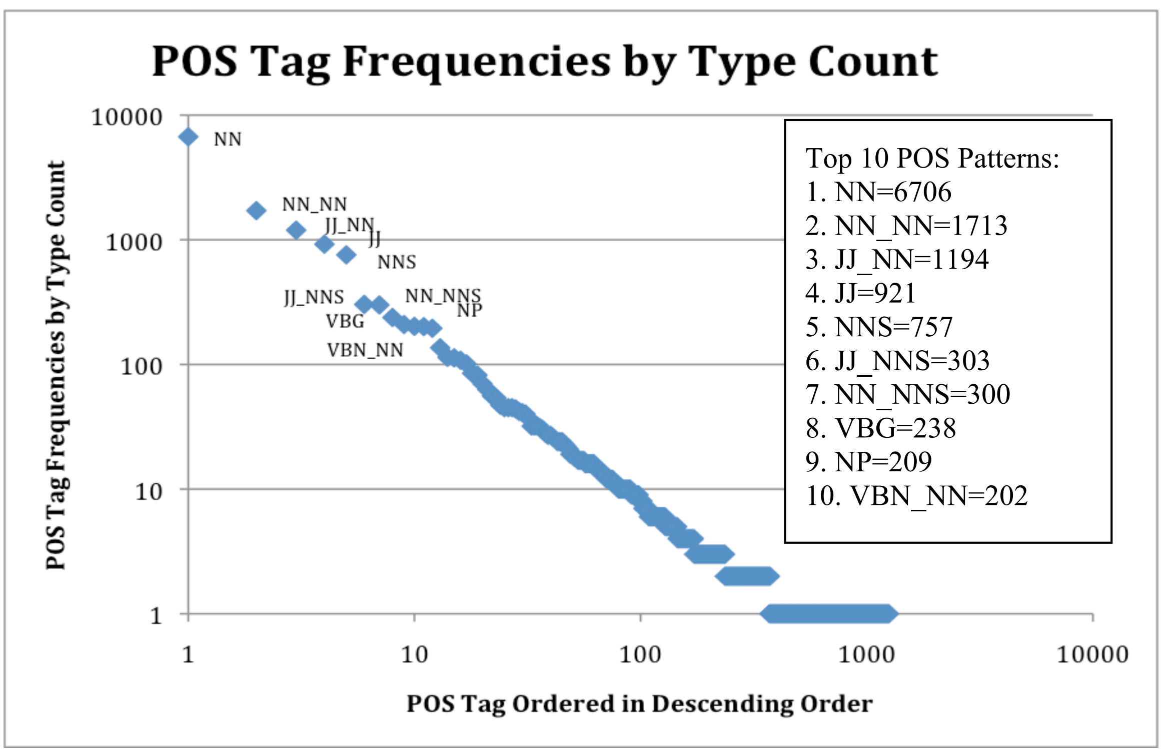

Figure 3. Part of Speech (POS) tag frequencies by type count.
Note that each axis is on a logarithmic scale.

(unexpectedly) adjectives with plural nouns are the next three categories in frequency.

The graph shows that the frequency of the POS patterns for tags follows a power law (Zipf, 1949); in other words, the frequencies of the POS patterns decrease exponentially so that a POS pattern is inversely proportional to its position in the list.

If one looks at POS patterns in the tag set based on type count (Figure 3) rather than token, the power law is still visible. However, we are reducing the frequency of occurrence of each POS pattern because we are only considering unique tags. This can be seen in Table 1 which presents the top 6 POS patterns ranked by frequency calculated by token and ranked by frequency calculated by type. In this table, one-word length POS patterns exhibited a significant reduction in frequency while the POS patterns of more than one word were not reduced as much. The two word POS pattern "NN_NN" and "JJ_NN" climb up the rank by getting a smaller reduction in frequency compared to "JJ" and "NNS". "NN" also exhibits a reduction in frequency but it is so frequent that it remains in the top rank.

Rank	POS Tag	Freq. by Token	POS Tag	Freq. by Type
1	NN	25205	NN	6706
2	JJ	6319	NN_NN	1713
3	NNS	4041	JJ_NN	1194
4	NN_NN	2257	JJ	921
5	JJ_NN	1792	NNS	757
6	VBG	1043	JJ_NNS	303

Table 1. The top 6 POS patterns ranked by frequency by token and frequency by type.

This can be interpreted to mean that, as expected, normalization conflates smaller word length tags more than bigger length tags. Overall token frequency of POS patterns is combination of the frequency of tags that can be conflated together plus the tags that cannot be conflated together and are different but have the same POS pattern. These two factors contribute differently to the frequency by tokens. For this reason, when we remove the tags that can be conflated, we observe a change in the ranking.

Part of speech tagging is integral to most NLP pipelines, since this step is a precursor to parsing. However, for social tags, parsing is not a meaningful step. Therefore, by studying the POS properties of tagsets in and of themselves, there is an opportunity to understand the nature of this kind of descriptive tagging. Linking POS data with other lexical resource information, and with semantic information may contribute ultimately to a deeper understanding of the nature of social tagging as linguistic data, and to the utilization of these tags in the museum context. The leap between using tags for access and understanding tags as a set of linguistic entities is the purpose of this research, so we are addressing relevant parts of this question in this paper

3.4 Theory-Driven Semantic Disambiguation by Domain

The second novel contribution of this paper is in the semantic disambiguation of tags by theory-driven distinctions. Identification of the subject matter expressed through social tags can provide an additional tool to understand, and thus control and manage, the noise created through the collection of this type of unstructured information. LaPlante et al. (n.d.) is undertaking a study to examine a set of 100 images of two-dimensional, representational paintings with 2909 unique tags in this specific collection.

While there are many theoretical approaches to categorizing the way an image can be described, from identifying a broad range of attributes (Jorgensen, 1998) to showing a hierarchical structure with levels of generality (Rorissa, 2008), there is still no consensus on the best approach to use (Stewart, 2010). To address this challenge, we are using a two-dimensional matrix based on the work of Shatford (1986) that reflects both the depth and breadth of information available about an image (Armitage and Enser, 1997; Bradley and Soergel, 2009). One axis of the matrix describes specificity, or an individual's depth of knowledge about the content of an image (generic (G), specific (S), and abstract (A)). The second describes the type of subject matter expressed (who (1), what (2), when (3), and where (4)). This core matrix was modified to include a visual elements category (V) as well as an unknown category (U) to capture information not related to subject matter.

Individuals from the museum community as well as project staff have categorized the tags assigned to these images using this two-dimensional matrix. Of the 2884 tags in the full collection where there was coder agreement, G1 (kind of person or thing) is the most frequently assigned category at 48%, followed by A2 (emotion or abstraction) at 10%, and G2 (kind of event, action, condition) at 10% (Table 2).

V: Visual	G1: Generic who	G2: Generic what	G3: Generic when	G4: Generic where
148 6%	1095 48%	227 10%	161 7%	32 1%
U: Unknown	**S1: Specific who**	**S2: Specific what**	**S3: Specific when**	**S4: Specific where**
216 9%	33 1%	5 0.2%	37 2%	62 3%
	A1	**A2**	**A3**	**A4**
	27 1%	236 10%	3 0.1%	2 0.1%

Table 2. Categorization of tags.

Looking at the specificity axis (Table 3), the overwhelming majority of the tags, 66%, are categorized as generic, followed by abstract at 12%.

V: Visual Elements	G: Generic	S: Specific	A: Abstract
148	1515	137	268
6%	66%	6%	12%

Table 3. Specificity of tags.

In Table 3, percentages are calculated against a total of 2284 tags. The 216 tags not include in this chart are those that were categorized as U (Unknown), which does not have a specificity component.

Table 4 shows that, on the type of subject matter axis (Table 4), the element most frequently expressed in the tag set (51%) is the "who" of the image, specifically, the person or thing represented in the image. The "what" of the image, or the event, action, condition, or emotion expressed in the image, was second at 20%.

1: Who?	2: What?	3: Where?	4: When?
1155	468	201	96
51%	20%	9%	4%

Table 4. Subject matter expressed by tags.

In Table 4, percentages are calculated against a total of 2284 tags. The 364 tags not included in this chart are those that were categorized as V (Visual Element) or U (Unknown), which do not have a subject matter component.

These figures were then compared to information gathered through other studies on a variety of different user groups and images collection. Overall, there are few similarities found between the types of tags assigned to images of art objects and those assigned to other image collections, showing that tag assignment may be domain-specific. It may also reflect Golder and Huberman's (2006) finding that a significant amount of tagging is done for personal use rather than public benefit, so the nature of the tagging task my impact tag type.

The importance of this analysis is that knowledge of this type of information can assist with managing the volume of unstructured tag information provided by users. It can help weigh the likelihood of different parts of speech in a tag set thus providing help in disambiguation. For example, the preponderance of tags expressing the who of an image would suggest that tags that are ambiguous such as gerunds are more likely to be nouns. It can also help visualize the type of information found in a tag set associated with art objects. For instance, this tag set can provide a substantial amount of generic information on things or events, but little valuable data on specific periods of time. Further research on this topic is being explored and will be reported in future work.

3.5 Original Contributions

Our overall research program addressed three questions, stated in Section 1. The novel contributions of this paper cross-cut these three questions. We have shown:

- Basic computational linguistic processing can impact tag analysis by token and type which will in turn affect down-stream tag analysis;
- Morphological and part of speech analysis impacts how tags are clustered and viewed;
- Computational linguistic tools can reduce some of the "noise" in tagsets;
- Theory-driven semantic analysis of tags reveals categories useful for disambiguation.

3.6 Future Work

Our future work addresses other aspects of the research questions set out in Section 1. As in Agichtein et al. (2008), we will be combining high quality content from museum sites with social tags. We will use the output of a toolkit (Anonymous) to identify named entities and noun phrases in texts associated with these images, provided by museum partners. Mapping information from existing text resources along with social tags raises challenges in concept relationships, disambiguation and then in sifting and filtering to improve object access.

Hsu and Chen (2008) examine tag normalization with respect to noun-noun compounds and their syntax. They utilize a spreading activation approach to normalize tags such as "drag and drop, draganddrop, dragndrop" to a canonical form based on a small manually created training set. The problem of these nouns in English is one we have not yet addressed but which is an important step in the language processing pipeline, especially for the handling the noise in social media. Edited text generally has guidelines for quality, whereas tagging does not.

It would also be valuable to analyze the temporal order of tagging based on user session to see if any patterns arise when looking at an individual user or at an individual session. For example, if in a given tagging session, a user tags one image with the words "red", "purple", and "green", can we use that information to disambiguate a less clear tag such as "gold" which could refer to either a color or a metal? Similarly, if we know that users tend to tag with nouns first, can we use that information to disambiguate tags in other tagging sessions?

Related research (Eleta, 2011) is addressing the issue of multilingual tagging and conceptual correspondences between languages and cultures.

In addition to these more general questions, there are some domain-specific questions that would be valuable to examine to help cultural heritage organizations manage large collections of tags. For instance, are there distinctions between the linguistic characteristics of tags provided based on object type, such as paintings or photographs? Similarly, are there distinctions between two- and three- dimensional objects or abstract and representational works of art? Based on initial observations, it appears that there are many lexical properties of tags that could be inferred from using information about object type, but this hypothesis is yet to be confirmed.

Acknowledgements

We wish to thank the members of the Museums Working Group of the T3: Text, Tags, and Trust project for their guidance and for their assistance in categorizing the tags used in this research. A full list of the member organizations may be found at http://t3.umiacs.umd.edu. We also thank Susan Chun, Independent Museum Consultant, for pushing our thinking on the question of linguistic processing of tags as well as Dr. Jennifer Golbeck and Irene Eleta, Ph.D. candidate in the University of Maryland College of Information Studies, for discussing the processing steps with us. Lastly, we thank the Institute of Museum and Library Services (IMLS) for their generous grant support of this project. Any views, findings, conclusions or recommendations expressed in this paper do not necessarily represent those of the IMLS.

References

Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, and Gilad Mishne. (2008). Finding high-quality content in social media. In *Proceedings of the international conference on Web search and web data mining* (WSDM '08). ACM, New York, NY, USA, 183-194. DOI=10.1145/1341531.1341557

Linda H. Armitage and Peter G.B. Enser. (1997). Analysis of user need in image archives. *Journal of Information Science, 23*(4). p. 287-299. DOI 10.1177/016555159702300403

Melanie Aurnhammer, Peter Hanappe, and Luc Steels. (2006). Integrating collaborative tagging and emergent semantics for image retrieval. In *WWW2006: Proceedings of the Collaborative Web Tagging Workshop*, Edinburgh, UK.

Hila Becker, Mor Naaman, and Luis Gravano. 2010. Learning similarity metrics for event identification in social media. In *Proceedings of the third ACM international conference on Web search and data mining* (WSDM '10). ACM, New York, NY, 291-300. DOI=10.1145/1718487.1718524

Grigory Begelman, Phillip Keller, and Frank Smadja. (2006). Automated tag clustering: Improving search and exploration in the tag space. In *WWW2006: Proceedings of the Collaborative Web Tagging Workshop*, Edinburgh. UK.

Steven Bird, Ewan Klein, and Edward Loper. (2009). *Natural language processing with Python: Analyzing text with the natural language toolset*. Sebastopol, CA: O'Reilly Media.

Rachael Bradley and Dagobert Soergel. (2009). *Image subject searching: What we know and where we need to go.* Unpublished manuscript.

Simon Corston-Oliver and Michael Gamon. (2004). Normalizing German and English inflectional morphology to improve statistical word alignment. In *Proceedings of the Conference of the Association for Machine Translation in the Americas*, Robert E. Frederking and Kathryn Taylor (Eds.) Washington, DC. 48—57

Irene Eleta. (2011). *Art images and multilingual social tagging: A museum without borders*. Computer Science Report No. CS-TR-4981, University of Maryland Institute for Advanced Computer Studies Report No. UMIACS-TR-2011-08, Language and Media Processing Laboratory Report No. LAMP-TR-156.

Scott A. Golder and Bernardo A. Huberman. (2006) The Structure of Collaborative Tagging Systems. *Journal of Information Science*, (32), p. 198-208

Ming-Hung Hsu and Hsin-Hsi Chen. (2008). Tag Normalization and Prediction for Effective Social Media Retrieval. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Volume 01 (WI-IAT '08).* IEEE Computer Society, Washington, DC, p. 770-774. DOI=10.1109/WIIAT.2008.92

J. Paul Getty Trust. (1988-2010a). *Art & Architecture Thesaurus (AAT).* Retrieved from www.getty.edu/research/conducting_research/vocabularies/aat/.

Corinne Jörgensen. (1998). Attributes of images in describing tasks. *Information Processing & Management, (34)*2-3, p. 161-174

Judith Klavans, Robert Stein, Susan Chun, and Raul Guerra. (2011). Computational linguistics in museums: Applications for cultural datasets. In *Museums and the Web 2011: Proceedings* (Philadelphia, PA), Jennifer Trant and David Bearman (Eds.).

Rebecca LaPlante, Judith Klavans, and Jennifer Golbeck. (n.d.). *Subject matter categorization of tags applied to images of art objects*. In progress.

Danielle H. Lee and Titus Schleyer. A comparison of meSH terms and CiteULike social tags as metadata for the same items. In *Proceedings of the 1st ACM International Health Informatics Symposium (IHI '10),* Tiffany Veinot (Ed.). ACM, New York, NY, p. 445-448. DOI=10.1145/1882992.1883060

George A. Miller. (1995). WordNet: a lexical database for English. *Communications of the ACM*, *38*(11), p. 39–41.

Bo Pang and Lillian Lee. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1-2), p. 1-135. http://dx.doi.org/10.1561/1500000011

Abebe Rorissa. (2008). User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory. *Information Processing & Management, 44*(5), 1741-1753. DOI: 10.1016/j.ipm.2008.03.004.

Sara Shatford. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly, 6*(3), p. 39-62.

Brian Stewart. (2010). Getting the picture: An exploratory study of current indexing practices in providing subject access to historic photographs / Se faire une image: une exploration des pratiques d'indexation courantes dans la fourniture de l'accès par thème à des photographies historiques. *Canadian Journal of Information and Library Science, 34*(3) (2010): 297-327. Retrieved from Project Muse at http://muse.jhu.edu/journals/canadian_journal_of_information_and_library_science/summary/v034/34.3.stewart.html.

Antonio Torralba, Rob Fergus, and William T. Freeman. (2008). 80 million tiny images: a large dataset for non-parametric object and scene recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*(11), p. 1958-1970. DOI=10.1109/TPAMI.2008.128.

Jennifer Trant. (2009). *Tagging, Folksonomy, and Art Museums: Results of steve.museum's research*. Available from http://conference.archimuse.com/blog/jtrant/stevemuseum_research_report_available_tagging_fo.

Jennifer Trant, David Bearman, and Susan Chun. (2007). The eye of the beholder: steve.museum and social tagging of museum collections. In *Proceedings of the International Cultural Heritage Informatics Meeting - ICHIM07* (Toronto, Canada).

Jennifer Trant and Bruce Wyman. (2006). Investigating social tagging and folksonomy in art museums with steve.museum. In *WWW2006: Proceedings of the Collaborative Web Tagging Workshop*, Edinburgh. UK.

Luis von Ahn, Ruoran Liu, and Manuel Blum. (2006).

Peekaboom: a game for locating objects in images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, p. 55-64. DOI=10.1145/1124772.1124782.

C. .J. van Rijsbergen, S.E. Robertson, and M.F. Porter, 1980. *New models in probabilistic information retrieval.* London: British Library. (British Library Research and Development Report, no. 5587).

Carol Vogel. (16 March, 2011). Four to follow. *The New York Times*, p. F24. Retrieved on March 29, 2011. Retrieved from http://www.nytimes.com/2011/03/17/arts/design/four-innovating-for-museums-online.html.

George K. Zipf.. (1949). *Human behavior and the principle of least effort*. Cambridge, Mass.: Addison-Wesley Press.