Abstract. Digital photography created a great need for storage and organization of photographs. One of the most popular organization methods used nowadays is tagging. It consists on classifying resources based on unstructured words and expressions. This allows for great flexibility and adaptability, but also brings issues. Having no semantic context, tags suffer from linguistic problems, which can hinder the retrieval of resources. In response, new methods of classification have appeared, one of them being narrative tagging, where users tell stories about the resources to classify. These narratives are represented as multiple separate elements with specific meaning. These elements, having a semantic context by design, do not suffer from the linguistic problems of simple tagging. This work follows up on a previous one that compared narrative tagging to regular tagging, when applied to photographs. It showed the advantages of narrative tagging, but left unanswered questions. In this work we try to answer them. An experiment was conducted where users uploaded 5 or more photographs and told a story about each of them. Each user also answered a survey on the usability of the system. By analyzing the collected data, we suggest changes to both narrative process and the interface of narrative-based systems, so that they become more appealing and straightforward to users, in order to increase this method’s acceptance and visibility.

1 Introduction

Digital photography led to a steep rise in the number of photographs a single person could take and keep. This amount of information, obviously, requires an organization system enabling people to store their photographs and reliably find them when needed. One of the simplest organization methods is to store them in a hierarchical directory structure. However, this system does not allow for any flexibility when searching for photographs.

Tagging is a recent method to classify and organize resources. It is an alternative and improvement over more traditional classification methods, as hierarchical taxonomies. It consists of simple expressions (tags) as classification
terms. There is no hierarchy or structure. This allows for great flexibility, but also to issues. Tags, not having any semantic context, are subject to several linguistic problems. These issues are especially noticed when searching for tagged resources.

One new form of resource classification and retrieving is called narrative tagging [8]. This method takes advantage of the natural tendency of people for telling stories, and asks users to create narratives about the resources being classified. These stories are told as a set of separate elements through a sequence of dialogs simulating a conversation. This approach has several advantages over tagging. First, users do not have to think of simple expressions without other context. Second, when retrieving, it is much easier to remember specific details of the needed resources than *ad-hoc* words chosen at the moment of their assignment.

A recent work done by Nuno Tomás [13] compared the narrative tagging approach to regular tagging in the context of photograph organization. The results of Tomás’ work showed that narrative tagging prompts the insertion of more information than regular tagging. Narrative tagging was also shown to allow for faster and more effective retrieval. Furthermore, the elements inserted in the narrative part of the application were shown not to suffer from the most important linguistic problems that regular tags do.

However, there were still unanswered questions left by Tomás’ work. No information was collected about the users’ opinion about the narrative system, or its perceived usability. Moreover, we have no data about the actual content of the stories. In this work we will try to address these problems. We conducted an experiment where users were requested to upload photographs and tell narratives about each one of them. We will analyze the inserted elements in detail to understand how the users tell their stories, and what can be changed in the story telling process to make it more efficient. We have also obtained the opinion of the users about the usability of the system.

At the end of this work, we will present important changes to the design of narrative-based applications, at both the user interface and the narrative process levels, to ensure that this kind of system is able to be accepted and adopted by users, so that narrative tagging is seen as a valid and desirable alternative for those that seek an organization method for their photographs.

2 Related work

In this section we will describe existing work in the area of tagging, with its advantages and disadvantages. We will also focus on image and multimedia tagging, since it is our area of interest. Lastly, we will study the existing research on narrative tagging, which will form the basis for our own work.

2.1 Advantages of and motivations for tagging

Tagging is an alternative to, and a paradigm shift from, the traditional process of rigidly and methodically classifying resources into taxonomies. When us-
ing tags to classify content, users are forming a folksonomy in a bottom-up fashion.

Shirky compares tagging to ontology-based classification, and presents tagging as a superior method of classification. For Shirky, the main advantage of tagging over ontology is the probabilistic approach on classification, allowing for a classification scheme adapted to the needs of each user, but can be used by all.

Choi compares tagging and expert-based classification. By comparing samples of web documents indexed by a website’s users to the same web documents indexed by professionals, he observed a greater consistency among website users than between the professional indexing experts. These findings assure tagging to be as reliable as professional indexing services, or more.

2.2 Tagging issues. Solutions and approaches

Despite of its advantages, tagging still suffers from problems. Tags, having no extra context, are subject to several linguistic problems. This means that using tags to classify resources can lead to difficulties when searching for and retrieving the tagged resources, or when collecting data from tag sets.

Andrews, Pane and Zaihrayeu recognize the main problems of tagging as a) base form variation (tags can be based on different forms of the same word), b) homography (one word can mean several distinct concepts), c) synonymy (different words have the same meaning) d) specificity gap (different users use words with disparate specificity to tag resources related to similar subjects). These problems can result in less accurate results when searching for resources based on tags. To mitigate them, the authors propose a new semantic model of representing folksonomies, creating many-to-many relationships between tags and concepts. The results reveal that linking tags to concepts fully automatically is inadequate, leading to very little matches. After manual preprocessing of the tags, however, they achieve a significant reduction on the tag vocabulary. They conclude that linking tags to concepts is effective, but only if tags are mapped to more than one concept, and not without previous manual intervention to “clean” the tag set.

A study attempting to remove ambiguity from tags, using both the context in which the tags are inserted and an ontology as basis for their meaning, is described by GarcÃa-Silva, Szomszor, Alani and Corcho. They use a semantically structured term repository as a dictionary to link tags to concepts. Tags, users, resources and the relation among them are represented as vectors, and the most adequate meanings for the tags are calculated through a similarity measure between the tags and the repository’s words. Their results are favorable, obtaining the correct meaning in the large majority of cases, showing a positive effect of taking advantage of the context when disambiguating tags.

An issue that is not usually as visible, in the medium to long term, is equally important or more, is that of loss of vocabulary efficiency. Chi and Mytkowicz investigate this by studying a data set extracted using information theory metrics, namely entropy, conditional entropy and mutual information. The authors
noticed that the entropy of the document set always increases, meaning that users continuously added new documents. In the tagging phase the tag distribution entropy was shown rising sharply until a point where it virtually stopped. The total tag number increases, which suggests that almost no new tags were created. This shows that users prefer using popular tags than creating new ones. The retrieval phase’s data shows that the entropy of documents conditional on tags increases rapidly, meaning that the entropy for all the documents referenced by each tag is increasing. This leads to tags being increasingly difficult to use as navigational aids, because when more documents are referenced, each one becomes harder to find. The authors also discover that documents and tags quickly become more independent of each other, and thus the usefulness of tags diminishes. They conclude that navigability becomes harder over time, even with a tagging system. They suggest that the most effective way of combating this problem is for the users to create more tags with which to describe documents.

2.3 Image and multimedia tagging

Multimedia items, especially photographs, are probably the most popular type of resource in the internet, being often classified by tagging.

Unfortunately, especially when images are concerned, users do not tend to invest much effort in assigning tags to resources. Shneiderman, Bederson and Drucker [12] state that most users spend their time mainly exploring photographs, and not adding information to them. The authors state that leading users to add more information should be a priority. They also say that the ability to share photographs is an important aspect, since users are more willing to tag photographs when they intend to share them.

Cao, Luo and Huang [2] focused on automatically assigning tags to photographs. Their approach is based on the fact that photos taken in the same context share several information and low-level features (e.g., location, date). Their approach is fully automatic, not requiring an initial set of tags to be supplied but is, however, restricted to a small set of tags. They developed a set of automatic classifiers to assign tags to the photographs which, by detecting similarities among images, through both metadata (e.g., time and date, GPS coordinates) and low-level features (e.g., color histogram), calculate the probability of each tag being indicated to each photograph. Results showed that this method performs better than simple direct automatic classification.

2.4 Narrative tagging

Gonçalves and Jorge [7,9] proposed a classification system where people tell stories about the resources. They studied stories told by volunteers, as well as data recalled by the users about personal documents, and constructed an archetypal story for this system. Based on that, they constructed two prototypes and validated them with further user tests, obtaining the results that prove that applications with the purpose of telling stories about applications can be built and will produce stories similar to those told to humans.
In a later work, Tomás [13] applies the narrative approach to the domain of photography organization. He compares narrative tagging to regular tagging. The results show that narratives, while taking more time than tags, allow users to insert much more information. For retrieval, the narrative approach also had better results. Finally, the story elements were shown not having any of the linguistic problems found in tags. Tomás concludes that narratives are more fit for classifying photographs than tags.

However, there is still work to be done in the topic of narratives as means of photograph classification. In our work, we intend to study the opinion of users about the narrative-based approach, as well as the contents of the stories to understand the suitability of the element types.

3 Implementation

In this section we will describe the implementation of our system, Narrative Based Image Repository (NarBIR). This system follows up on Tomás’ [13] study, maintaining a similar interface and process, but introducing changes in order to obtain better and finer results.

3.1 Interface

Here we will describe the graphical user interface of NarBIR. We will explain how the different stages of the NarBIR workflow were designed and the reasons behind them.

Photograph uploading and narrative creation

We shall start with the photograph uploading and story telling process. Users were asked to upload a minimum of 5 photographs.

The upload is performed through a page where users may either drag their files and drop them on to the page or press a button to open a file chooser dialog. Multiple files may be uploaded at once, although only image files are accepted. When the files are dropped or chosen, the upload begins, illustrated by a graphical progress bar. Each file's upload may be stopped through the Cancel button. After being uploaded, the files may be deleted from the server with the Delete button. When finished, the user may press one of the buttons marked “I'm done! On to the next step!” to continue.

After the upload process is the narrative screen. On the left side the current photograph is presented. On the right side is the narrative area, where the story is told. This layout attempts to maximize the available space for the story, while maintaining a clear view of the photograph.

The narrative, like in Tomás’ work [13], simulates a conversation between the user and the system, to stimulate the users’ memory. Each of these dialogs
request information about a certain type of element. The possible element types are:

- a) **Place**, b) **Time**, c) **Device**, d) **Quality**, e) **Dimension**, f) **Author**, g) **People**, h) **Description**, i) **Purpose**, j) **Orientation**, and k) **Event**. There are two special values that can be assigned to every element. I **don’t know**, for when the user cannot recall the correct value, and N/A, for when there is no applicable value. Users may also skip the current element and go to the next in order, through the **Tell later** link, or skip to a specific element using the **Tell about...** link.

Users may insert values of the same type multiple times in each narrative, in order to achieve the desired level of detail (e.g. writing both the names of the city and the country in the **Place** element). To ensure that no elements are lost in any case, each element is saved immediately after its value is set, through AJAX requests to the server.

The elements in the narrative are represented by links which, when clicked, allow to edit the element’s value. When being edited, an element may also be deleted through the **Delete** link, which is normally disabled. When the user considers the narrative finished he may press the **OK. Next one!** button to go to the next photograph without a story. When all of them are narrated, that button’s text changes to **All Done!**.

The final step is to review the narratives. In the review page users can view all of their photographs and the stories they told about them. In the left side there is a list with the photographs’ thumbnails. The user may click on them to load the story for each photograph. In the right side the photograph is shown on top, while below is its related narrative. Users may press the **Edit this story** button to be taken again to the narrative page and edit the narrative.

**Surveys** We developed a small survey to collect basic details of the users’ identity. This survey’s questions are about

- a) **gender**, b) **age**, c) **academic qualifications**, d) **current employment**, e) **type of computer usage**, f) **frequency of computer usage**, g) **frequency of photography**, h) **ownership of camera phone**, i) **ownership of regular camera**, j) **profession as a photographer**, k) **most used camera type**, and l) **preferred usage of analog or digital camera**.

We also intend to understand if NarBIR’s workflow and interface, as well as the narrative process, are seen as usable and useful by the users. For that purpose we developed a second survey, dubbed **Debriefing**, which the users answer after telling all the stories. Besides questions defined specifically for the needs of this study, this survey includes the System Usability Scale (SUS) [1] questionnaire (with a slight change [2] to aid non-native English speakers), for a standard comparable measure of usability

In this survey the users were asked (not including the SUS questions) a) how they would classify the amount of work needed, b) how they would classify the amount of information inserted, c) if they would consider using this system to classify photographs, d) if they have ever used tagging in a site or application,
e) if the prefer tagging or the narrative approach, f) to state any negative aspects of the application, g) to write any suggestions they thought useful, h) to write any final comments.

3.2 Database

In NarBIR’s data model, the entities we store are a) Users, b) Photographs, c) Stories, d) Survey Answers, for the first survey, e) Debriefing Answers, for the final survey, and f) Instrumentation Events, capturing the actions of the users.

3.3 Technologies and programming languages

Our system is a typical web application, with PHP: Hypertext Preprocessor (PHP)\(^1\) version 5.3.3-7, as the server-side component, HyperText Markup Language (HTML) and JavaScript for the client side, using Asynchronous Javascript And XML (AJAX) for asynchronous communication, with JavaScript Object Notation (JSON) instead of eXtensible Markup Language (XML). The database uses MySQL\(^2\) version 14.14. All of the scripts and stored procedures are written in Structured Query Language (SQL).

4 Results

In this section we will present the results obtained during the usage of NarBIR, and compare them to Tomás’\(^3\) to understand what has changed.

Statistics and comparison to previous work

On average our users took more time to tell their narratives (3 minutes and 31 seconds, versus 2 minutes and 2 seconds from Tomás’ results), but they have also inserted more elements per narrative (an average of 10.36 elements versus 7.96).

The difference in time may be due to differences between the two interfaces, or to Tomás’ users having inserted more stories (an average of 32.8 versus 5.27 from our users), thus having more experience. The difference in number may be due to our dialog placement (along the story element instead of fixed) being less tiresome, conserving the users’ patience. We noticed that users showed a tendency to follow the default order of elements, not frequently skipping them or changing the order of the narratives.

In the debriefing survey, the answers showed that most users found our system useful and would consider using a narrative-based approach in the future. The average SUS score obtained of all the respective answers was of 78.13, which is very positive.

\(1\) http://php.net
\(2\) http://www.mysql.com
Group narratives We asked two groups of users to each tell stories about a set of photographs, all taken in the same event (one different event and set of photographs per group), in which all of the members participated. We analyzed the values of the narratives’ elements, to understand if the values are equal, can at least be deemed the same by a person (if not by a machine), or are simply different.

The Purpose element was the least compatible, having different values between users in nearly every narrative. The Description element had mostly compatible values, with the exceptions being due to the use of personal comments. The Time element was always consistent among narratives and its values, while not always compatible, can be seen as complementary to one another. The Dimension and Quality elements were mainly equal, except that in nearly every photograph one user’s value always diverged from the others. This may be due to the subjectivity of the possible values for these elements. The Orientation element, showed equal values in all photographs except for one, where the user kept the same value inserted in all other narratives, perhaps by habit. In the Device element, at least two out of four users agreed in all narratives. The remaining users inserted only the camera model or “I don’t know”. The Place, Author, People and Event elements were mostly compatible among all photographs, with more or less details being inserted by different users.

Element value analysis Analyzing in detail the values of the elements inserted in all the narratives, we noticed some particularities. Users were confirmed to use values of very disparate granularities, especially in the Device element. In the Place element, some users included several values in a single element (e.g. “Rome, Italy”) instead of using multiple elements. The name of the Event was substantially used as values for the Place and Time elements. In the People element users inserted references to animals and objects. The Description, Purpose and Event elements had the common property that users tended to insert values apt to belong in other element types.

4.1 Lessons and improvements

Here we shall describe what we have learned and suggest changes to future narrative-based systems, in order to improve the users’ experience.

Usability We have learned that the usability of these systems must be a priority. Users must be allowed to tell stories about multiple photographs, in order not to repeat information. Afterwards, it should be possible to specify the information in each photograph with more narratives. The steps and details of the narrative process must be clarified to avoid confusion. This is especially true of the possibility to insert multiple elements of each type.

Suggestion of values and auto-completion for the story elements should be implemented, in order to reduce the effort needed. However, measures should be taken not to expose delicate information. Similar photos should be detected and suitable elements suggested to be propagated among their narratives, avoiding
repetition. The story elements should be able to be reorganized, both automatically (change the default element order) and by the user. Navigation through other narratives and photographs should be possible while telling a story, in order to confirm any detail, improving consistency.

Information extraction To save the users some work, information should be automatically extracted from photography metadata (e.g. date, location) and added as story elements, avoiding manual insertion of those values. Elements’ values should also be analyzed and any suitable sub-elements extracted and added as elements to the story.

Narrative process The only possible new element we see is Comments, in order to keep the personal comments, leaving Description for the visual descriptions.

A social aspect should be introduced, perhaps by integration with other social networks, to allow connecting to friends and sharing photographs and stories. Users should be allowed to tell stories about each others; photographs, to obtain more information. In this case, all of each picture’s stories should be shared to improve consistency.

Finally, users have also asked to be able to add new element types ad-hoc, and to change the way in which the narratives are told (e.g. adding different emphasis and verb tenses to elements). We believe these would be very interesting changes, but must be carefully studied first.

5 Conclusion

Tagging appeared recently as an innovative approach to classification. It brings flexibility and adaptability but, because tags have no associated semantic context, suffer from linguistic issues. Among several methods created to mitigate these issues, we are especially interested in a new approach to classification, narrative tagging. In this approach, instead of using unstructured expressions as information, users are asked to tell the story of each resource.

These narratives are composed of several distinct elements. Their values are automatically included in specific contexts and do not suffer from the linguistic problems of tags. Narrative-based systems also compel users to insert more information than in equivalent tagging systems.

A recent work studied the use of narratives for the organization of photographs. We have taken this work as a basis and built upon it to understand what information is inserted into the narratives. For that we built a narrative-based system and used it to collect photographs from users and ask them to tell stories about each. In the end they answered a survey about their opinion of the system, including the SUS questions for added reliability.

We performed a case study where two groups of users told stories about the same sets of photographs (one for each group, all photographs taken in events where all of the group members participated) in order to study the compatibility of the several narratives, gathering more information for the development of the social aspect of narrative-based systems.
We collected less data than the previous work, but were able to obtain more data in each narrative. In the group data analysis we saw some inconsistency, but in general the stories were compatible. The opinions about usability were positive, as well as the SUS score. From the free-form questions of the final survey we noticed some important issues and suggested changes to mitigate them in future systems.

The majority of the users seem to be receptive to the narrative approach and find the story telling process to be usable and not requiring excessive work. If some of our suggestions are confirmed to improve the usability of these systems and implemented, it is very likely that narrative classification systems will soon be more well-known and widespread, allowing for much more effective photograph organization systems.

References

1. J. Brooke. SUS: A Quick and Dirty Usability Scale.
4. Y. Choi. Usefulness of social tagging in organizing and providing access to the web: An analysis of indexing consistency and quality. 2011.
5. K. Finstad. The system usability scale and non-native English speakers.