A Narrative-Based Structured Alternative to Tagging
Nuno Tomás, Tiago Guerreiro, Joaquim A. Jorge, Daniel Gonçalves
INESC-ID/DEI, IST
Av. Rovisco Pais, 1000 Lisboa
nmrt@ist.utl.pt, tjvg@vimmi.inesc-id.pt, {jaj,daniel.goncalves}@inesc-id.pt

ABSTRACT
The increase of multimedia information over the past few years has led to the creation of mechanisms that enable its organization through descriptions or keywords. A popular way of associating such descriptions to content is tagging. It can be found in popular sites such as Flickr (for images) or Delicious (bookmarks), among many others. The user is free to associate several tags to content, richly describing that content, which may lead to an easier retrieval at a later time. This solves several problems with hierarchical organization schemes, where choosing the “correct” classification is often hard. However, tagging is completely unstructured, leading to problems of its own. If cataloguing new items is easy, nothing guarantees that the tags used are most appropriate or the same tags are used in similar situations. Problems such as polysemy, synonyms or the problem of basic levels of language make the reuse of tags lower, and hinder the subsequent recovery of items.

We propose an approach in which narratives are used as an organizing principle for the tagging of photos. We are used to tell and hear stories, making them a natural form of interaction. By inter-relating the information in them as a coherent whole, stories convey much data in a structured way. By resorting to narratives describing their photos, the users are able to convey more information than by using traditional tagging methods. Furthermore this increases tag reuse and minimizes other problems found in tagging. We conducted a user study with 40 users over a period of three months, which shows this to be the case.

Author Keywords
Tagging, Narrative-based interfaces, Digital Photographs

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces. H.3.3 Information Search and Retrieval:

INTRODUCTION
The increasing access and use of several computing devices by the same user means that they increasingly need to find effective ways to organize and retrieve their personal information. Many users are resorting to online services to prevent this fragmentation, by storing their information online instead of in individual computers and other devices. Thus, the data needs only reside in one place, and can be readily accessible from everywhere there is a network connection. This is done, especially, for particular kinds of data, such as bookmarks (eg: delicious¹), documents (eg: Google Docs²) and digital photographs, where Flickr³ is probably the most popular choice.

The amount of information stored using those services leads to other problems. Namely, it is hard to retrieve objects at a later time. Some kind of classification scheme must be used to organize the information to facilitate its retrieval. Given the relatively informal and unstructured nature of the world-wide-web, and the fact that web applications run in the browser and, thus have limited access to the users’ computers and limited computational resources at their disposal, it was only natural that simple yet effective ways of organizing data would arise.

One of the most used and increasingly popular is tagging. This system allows users to freely associate tags to content. It is nowadays present in most sites, and is both easy and effective. It helps solve one of the greatest problems with traditional, hierarchy-based, classification schemes: the user can assign several tags to the same object. When only one category can be chosen (such as in the filesystem), great care must be taken when choosing the classification. With tagging, the users can associate all the tags they find relevant and that they feel might be of use in different, future, retrieval contexts.

Furthermore, tagging presents a collaborative aspect, since in most sites the users can search for content that is not their own, and sometimes tag such content. From this behaviour a folksonomy [7] might arise: a set of popular tags shared by a community, used by general consensus to describe certain

¹ http://delicious.com/
² http://docs.google.com/
³ http://www.flickr.com
objects. In essence, a folksonomy is a shared language understood by the members of a community.

Although very popular, these tagging systems suffer some problems [2]. One of the most relevant is *polysemy*, when the same word can have several meanings. This might lead to unexpected search results, as the system is unable to separate those meanings. *Synonyms* are also a problem because different words can have the same meaning. When doing the search the users must use all the synonyms used for the content they are looking for, so that the results are as complete. A third problem is *basic level variation*. The same item can be described in different ways, all correct, but describing it in different ways.

These problems lead to *low tag reuse*. While ideally the same tags would be used by everyone in a consistent way, in practice that is not the case. Not only do different people use different tags to describe the same objects, but the same user might change the tags in use as time goes by. A few tags, the more predominant in the folksonomy, are used recurrently, but most of them do not. This leads to a severe fragmentation of the classification scheme, with dire consequences when retrieving the tagged objects.

We propose that a semi-structured way to describe objects in general and digital photos in particular can help alleviate this problem. By allowing the users to craft narratives describing their photos, they will not only convey more information but also help them maintain more consistent descriptions.

Narratives are a more natural and structured approach [4]. Storytelling is something that is intrinsic to human nature, since we grew up listening to stories told by our parents and grandparents. Being so natural, the narratives become an assumption quite plausible to associate information. When telling a story, people will add key attributes that are associated with the resources, such as “what is it” or “to whom it belongs to?”. This way a person forms a structure that has a beginning, middle and an end, unlike tagging, in which the user must recall or “invent” all attributes associated to content, with little help, which might hinder the retrieval process. Taking this into account, we think through the use of narratives, we can significantly alleviate problems such as ambiguity, polysemy, and synonyms, which will result in a higher tag reuse.

To validate our approach, we started by carrying out a study on how users tell the stories about their photos. From that we found the most relevant elements in those stories, and how they are related in them, allowing us to infer patterns in their flow and how they are typically organized.

This study led us to consider a solution in which the user is guided throughout history, in a succession of dialogues corresponding to the different relevant story elements, as found in the previous study, following a policy of suggestive tagging.

In the second phase, we developed a website where users could store their personal photos and then associate one or more stories to each photo. An alternate version of the site resorting to traditional tagging was used by a different set of users. We measured their behaviours over the course of three months, and found, among other results, that indeed narrative-based tagging can lead to a tag reuse of 89%, vs 36% for traditional tagging approaches.

In Section 2, we give an overview of tagging systems and their major problems. Section 3 presents the study about photo-describing stories. The interface of the website created for the user studies is presented and explained in Section 4. In Section 5 we present the results from the user evaluation, after which we conclude presenting possible avenues for future work.

**RELATED WORK**

In recent years, with the astronomical rise in popularity of a community.

In Section 2, we give an overview of tagging systems and their major problems. Section 3 presents the study about photo-describing stories. The interface of the website created for the user studies is presented and explained in Section 4. In Section 5 we present the results from the user evaluation, after which we conclude presenting possible avenues for future work.

**RELATED WORK**

In recent years, with the astronomical rise information quantities common users must deal with on a daily basis, tagging quickly rose in popularity, as it enables users to describe their content very easily and provides a good basis for its later retrieval [6].

There are numerous web sites using tagging as basis for cataloguing and searching content. Examples include Flickr, Delicious, Youtube, and many others. There are different types of tagging, but in the context of this study, we only focus our attention on tagging systems where the user has the responsibility to associate tags to content.

We studied several sites’ different tagging strategies in depth. Two of the most relevant for our study were Flickr and delicious.

Flickr [8] is a website for sharing and hosting photos and is based on tags, which help to share, search and navigate the contributions of the entire community. Each user has the ability to store their personal photos adding their own tags to describe them. Tags have a crucial role in this site, being the main navigation and retrieval tool at the users’ disposal.

As for the process of tagging, only the content creators are by default directly empowered to add tags to that content. It is also possible to delegate rights to other users, which allows them also to add tags to other users’ content. At the time of this writing, when adding tags the system offers no suggestions (blind-tagging) of tags given to similar resources. Another design option followed by Flickr is that, in terms of interface, no repeated tags for the same resource as allowed.

Delicious offers a web service that allows users to add and search bookmarks. A motivation of this web site comes from the proliferation of online devices that each user has. Thus this website permits the centralization of bookmarks, so that they are accessible from any location. Also, it creates a bookmark sharing community, giving users in the community access to each other’s users bookmarks. Like Flickr, tagging is the primary way of organizing data [2] [7]. The rights of tagging on resources are available to
everyone: any user can associate a tag on a resource posted by another user. When new tags are being entered by a user, the system offers suggestions of possibly relevant tags, based on the classifications made by other members of the community of the site being bookmarked and related sites. The system accepts several tags for the same bookmark, which may result in duplicate tags from different users.

Studies of the aforementioned sites revealed several problems related to tagging [2] [7]. Polysemy that occurs when a word has multiple meanings and their meaning can only be perceived in context of the sentence where it appears. One example is the word “position”, which can be applied in various contexts such as: “Put yourself in a correct position for the picture” and "achieved a good position in the company". While in the first sentence word "position" refers to the person back position, in second sentence refers to the level reached in a company (administrator, manager, etc.). It is a problem because a search can return results that are not related to the meaning that is sought. Superficially, polysemy is similar to homonyms, where a word can have multiple meanings that are not related to each other’s. However, the homonyms are not a problem as complicated as polysemy, since they can be largely excluded by the adding additional tags to the search query, referring to related concepts. This will potentially filter out unwanted results. An example is the word “fluke”. It may refer to the fluke of an anchor or to a goal that happen from a fluke. In this example to solve the problem of homonyms when doing the search, we can add also the word “anchor”, thereby discarding any results that don’t have any similarity with what we are looking for.

Synonyms present another challenge because many words can have the same or very similar meaning. It is a complicated problem in a collaborative system due to the inconsistency of the terms used in tagging, making the task of searching truly complicated. The solution is, to have all users reach a consensus (often impossible), or have the notion that they have to enter multiple combinations of tags during searches in order to account for all possible outcomes.

Another problem associated with tagging is related to the basic levels of language. In fact, the specificity with which each user describes a piece of content can vary, depending on their experience or degree of specificity in the area. For example, an animal lover and particularly a dog specialist can associate a tag to a new picture, using the breed or any other feature of a dog, while for the average user it will be a photo of "dog". Both variations in the degree of knowledge, as changes in social and cultural from the users, may cause variations on the basic level.

There are also cognitive problems [5] associated to the fact that a user has to remember a certain tag associated to a content, which may be somehow complicated after some time.

All these problems reveal themselves in a low tag reuse. By being hard to decide which tags should be used when and with that meaning, lots of different tags proliferate in these systems. While some very common tags are indeed reused very often, that is not the case for the large majority of them. For instance, a study of tags used in delicious [2] has shown that most tags are used by a very small number of users, following a power-law. This is true both for a single user and for the entire community.

In research made about the theme, no system was found that used the narratives as a way help the user to associate information to the photos. With some variations, the existing systems are based mainly on simple tagging as a way to do so.

**STORIES ABOUT PHOTOS**

We must take into account that a story is not just a series of elements interconnected without apparent logic, but elements that, together, convey a consistent structure with well-defined rules. As we shall see, it is possible to identify certain elements that appear more or less often in the stories told by people. When designing the narrative-based interface we take into account not only each story element individually, but also the relationships between them and the overall story structure. This structure must be recreated so that the user can tell story as naturally as possible. This will ultimately provide us with stories that capture the most important elements describing the user’s photos.

To find out exactly what to expect in photo-describing stories and with which structure, we conducted a user study. We interviewed 20 people asking them to describe their photos. To each person was asked to tell three stories about three personal photos, which was then transcribed and analyzed yielding insights about the elements that made up the story, its structure, and the logical course it followed. The results were the basis for the construction of our test interface.

Part of the users was composed of heavy internet users, and previous users of photo hosting sites such as Flickr and hi5. Were also took care to included in the group people that with an older age, looking for possible different views and sensitivities regarding the subject of our research.

**Procedure**

Twenty interviews were conducted in order to get a good sample of possible stories. Each user told three different stories, thus giving us a total of 60 stories to analyse.

Each interview consisted of two parts. In a first phase was described to the interviewee, which was the objective of the work and what was expected of the interview, showing how their part in this work. It was also collected relevant information from interviewee to use in statistical data (age, occupation, gender, etc.). Was then asked the possibility of recording the interview, with privacy of personal data to be collected guaranteed. Request that all interviewees agreed.
In the second part of the interview, was asked that the interviewed, tell the story of a personal photo. We used a semi-directive interview to ensure that the interviewee is free to tell its history and the interviewer has the possibility to guide the interviewee, if he somehow divert from the objectives of the interview. The interviewer had at his disposal a number of guide questions for the different elements of the stories, which were intended, to collect information relevant and lead the interviewee to use this topic to continue his story. The questions were reduced to a minimum since the purpose wasn’t a question-answer dialogue, but mostly the respondent to be the "protagonist".

Our goal was to get stories that the users would naturally and normally create. Thus, the interviewer's role was reduced to a minimum, intervening only in situations where the interviewee was blocked, at a loss about what to say next. In those situations, a set of predefined sentences, designed to prompt the user for more information without suggesting any particular thing to them were used. To enable a subsequent interviews analysis, we used a tape recorder, after ensuring the recordings and transcripts would remain private and obtaining the users’ agreement before starting to record. This prevented undue pauses that might break the natural rhythm of the story, that might skew or invalidate de results of our analysis.

The subjects were asked to choose three personal photos, and then tell the story of each one. The only requirement was that the interviewee had to remember the photo, so that a story could be told. It would not be correct to show them a predetermined set of photos, as only about their personal pictures will they have a story to tell. Indeed, stories often contain autobiographic and contextual elements that may be important to a particular user, but not to others. For that reason, we dealt only with personal photographs.

As previously stated, the privacy of the contents of interviews was preserved by removing from the transcripts possible references to personal aspects in the photos’ descriptions that might be used to identify their creators.

Analyzing the interviews
In order to collect all the elements that made the stories of the users photos, these same stories were transcribed and were subjected to analysis content. This analysis was performed manually, since it is impossible to automate the identification of the elements due to their variability. Although keywords can be associated to the elements, they also can be composed of sequences of words and entire phrases, making it difficult to automatically classify and understand them. The analysis followed the guide, which we specify below, that describes what the possible story elements we looked for in the stories. This guide was used in each of the interviews in order to achieve a coherent analysis of the entire sample.

We conducted two preliminary interviews, before the ones actually used for the results of our study. They allowed us to fine-tune our procedure, and to validate the list of possible story elements. With the information collected from them it was possible to identify new elements that were not part of the list elements initially identified and also refine the set of questions in order to adapt better to the stories. they also served as a means preparation for the interviewer. The list of items with their questions remained open to new elements that would possibility appear after the analysis of the interviews, so that the final list would be the most complete as possible and that all the elements identified the stories occupy a place on the list. As it turns out, no new elements were found.

Below is an example of a story told by a user, which can confirm the great amount of information that is transmitted in a story:

"This first picture was taken in April this year, the wedding of my older cousin ... I’m in the photo, my sister and my grandmother also... ah, I am a little better dressed than usual, jacket, beige trousers. This picture was taken in Loures... in the afternoon. A lot of people were there. It was the marriage of cousins here in the father's side, was the marriage of the oldest daughter, that as 35 years. ”

The elements considered in the analysis of the stories were:
1. Time: Time references to when the photograph was taken, including specific holidays (Christmas, Easter, etc.) Or seasons (summer, Winter, etc.). Other timeframes should be classified as events (birthdays, days I was travelling, etc.).
2. Location: Location where the photo was taken whether geographic locations (Lisbon, Switzerland, etc.) or benchmarks (IST, Eifel Tower, etc.).
3. Author: the person who was responsible for taking the picture, may be a name (Mary, Joseph, etc.) or may be described by a degree of kinship, a friend or a passer-by;
4. Purpose: Reason claimed why the photo was taken (because it was a funny situation, to remember the great day, etc.).
5. Photo Type: The photo is a picture of a person or persons, being considered a portrait, or is an image that was intended only capture a landscape (cascade mountain, etc.) or a built environment such as monuments
6. Size: Reference to the size of the photo (large, medium, normal, small, etc.).
7. Event: Event of personal life that is in some way connected to the photo or the holder's photograph (trip to Italy, his grandfather's Birthday, etc.)
8. Device: Reference to a tech used to take a photo (machine digital / analogue, phone, etc.).
9. Description: content that describes the picture (my hand, sun surrounded by clouds, etc.), or expressions that are
being represented (we with a very cheerful smile, we have the flash in our eyes, etc.).

10. People: Reference to persons who are both in picture as those in the story behind the photo;

11. Quality: Reference to the photo quality (good, bad, etc.).

Information was collected about not only whether an element occurred in a story or not, but also about the frequency of the elements (some appeared more than once per story) to give us an ideal of which were the most easily remembered by users, and may take these as the most relevant.

We distinguished two types of elements: the induced, ie, those that were elicited by some intervention by the interviewer to help the interviewees to resume their stories, and the spontaneous were people didn’t needed any reminder and naturally reported the element. Negative responses to questions from the interviewer were also considered. For instance, mentioning "there were no more people in the group" is different from not knowing if there were more people or not. In the first case the subject knows something, in the second he does not. This distinction was also taken into account in the analysis of the interviews.

As mentioned above, besides making an analysis of occurrences of each element we also made a relational analysis. We took into consideration the number of times that a specific element precedes another, and that can reveal the order in which the elements have been referred, giving an idea about the relations between them. We account only for transitions where the second element wasn’t induced, as only then would an actual mental connection between the elements exist. Induced elements might bear no relation to the ones that came before, since there was an intervention by the researcher between them. This relational analysis, in essence, gives us an idea about the typical structures of photo-describing stories.

Results
This section analyzes the results of interviews.

Characteristics of the interviewees

![Figure 1 - Interviewees profile](image)

Twenty people were interviewed, ordered by their Internet experience. In particular it was important that they were users of websites such as Flickr and Hi5. Analyzing in detail, thirteen (65%) were males, and seven (35%) were female (Figure 1a), with ages between 18 to 57 years (Figure 1b). In relation to their studies, are mostly people who attended some kind of college undergraduate course. About 40% of the people work, and only one person has an academic degree, but is unemployed. Most have contact with a computer and are regular users of the Internet (90%), and many of these people have an account in websites that store personal photos (85%). There are two people who belong to a group of different age and have little or no contact with computers. These users were added to the group of interviewees to explore the possibility of more experienced people, with other life experience could tell stories with different content. We did not find any such differences. This group of people, although not very diversified in terms of age, is representative of most users of the services we are considering in our research.

Story Elements
In this section we present the elements that we recorded from the stories told by users for their personal photos. We can observe the frequency of elements in Figure 2. Elements such as "Location" and "Time" are the ones used with greater frequency and therefore, in general, are more easily remembered.

In Figure 2 are also represented the frequencies that the elements were induced in the stories. As can also check the "Author", "Device" "Type" and "Dimension" were the elements that most often were induced in the stories. But as we have seen, for example the element "Author" or "Device" are elements with high frequency, indicating that may be important elements but not easily remembered by the user, without having some type of help.

![Figure 2 - Elements frequency](image)
elements considered important as they are reminded easily and are mentioned most often spontaneously.

Something that has to be taken into account is that the frequency may not give a real perception of the importance of the element. For example the "Time" element was mentioned 82 times. In the number of times that it was present, if the element is distributed in only eighth of all 60 stories made, it’s meaning it’s not as relevant as if the references are distributed by 20 stories. Being spread over more stories may indicate that it is an element that is used repeatedly to tell the stories of photos, then more important when describing a photo.

Figure 3 shows the balance between the frequency of elements with the number of stories where the element was present. We can see that only the elements "Time" and "Local" have a frequency well above the number of occurrences. This may mean that, or the elements are complicated to describe, or users make several references in order to clarify the information. Even so, they were mentioned in almost all of the 60 stories analized.

![Figure 3 - Frequency vs Occurrence](image)

**Story Structure**

In a narrative the elements that constitute them, by themselves have a relative relevance. It is important also consider how they relate between them. In order to examine this aspect we performed a relational analysis, recording all transitions between the elements of stories. As mentioned above, only transitions between elements that arose spontaneously in history were considered. As was to be expected, we found a number of pairs in which there was no transition between the elements in the 60 stories that have been recorded. Given the memory association that is characteristic of humans, there are elements pairs that were not mentioned together, which may mean that people, when structuring stories, do not associate the two elements, not considering a relationship between them.

On the other hand, there are some pairs whose relationship is strong showing considerable frequency transitions. We can see three pairs of elements with transitions quite strong: “Time-Local”, “Local-Event” and “Device-Quality”. Other pairs to consider are: “Local-Time”, “People-Description”, “Location-People” and “Event-Time” (sorted by decreasing of importance). Something that could be a relevant factor and lead to wrong indicators is the fact that element pairs with large number of transitions between them, may be a result of being elements with high frequency. For example, if the elements A and B had a frequency of 4 and had 3 transitions between them, the value of transitions, 3, would be lower if the pair C and D had frequency of 10 but had 4 transitions between them. Although C and D present more transitions, the ratio of A and B is clearly stronger, where B follows A 75% of the time, D only follows C in 40% of the time. In order to get around this problem we have chosen to normalize the data, making the number of transitions divided by the sum of the spontaneous frequency elements concerning the transitions. We chose the frequency of spontaneous rather overall frequency, because the transitions were calculated from the elements that arose spontaneously in the interview. As observed, the transitions with stronger relationships are “Device-Quality”, “Time–Location” and “Location-Event”.

**Archetypical Stories**

As experienced in previous sections, the photo stories contain patterns, sets of properties that are present in most of the interviews, as exemplified by the transitions that were analyzed in the previous section. However, there isn’t an easy way to transfer them to the tagging interface narrative. It was therefore necessary to use our findings to infer an archetypical story structure that could be used as the basis for the design of the interface. To do so we took into account factors such as the frequency of elements the stories, the transition between the elements and their occurrence.

One-way to organize the story with all these factors is by using Hidden Markov Models [9], trained with the sequences of elements that we have extracted the interviews. The elements were not separated between induced and spontaneous as it may remove the general idea of the story.

The overall story is composed by 11 elements. The first element is what happened with more often in the first position in the stories used in training the algorithm. Is then used the transition probability between elements to be able to find the next element of the story. The general history is made in this order: Location, Time, Device, Quality, Size, Author, People, Description, Purpose, Type and Event.

**Design Guidelines**

The results presented in previous sections, serve to check and withdraw some basic principles for the interface design, and we will outline them in this section.

**Customization**

There were no relevant factors such as age or gender that affect on a large scale in how stories are told. Although some statistics, as the size of the stories or the ratio of transitions between elements or the size, show some results
of which we can draw some conclusions, do not seem so important, so that may reflect the need to create different interfaces for both genders or for different age groups.

**Dialogues**
A large number in the frequency of the elements of stories had to be induced by the interviewer. Some elements appeared more often than others as induced elements, but still appeared to be important to interviewees showing that they must indeed, whenever possible, be part of the stories, as they convey relevant information about a photo.

This shows the importance of dialogue with the users, allowing them to be able to get all possible information

**Elements and Structure**
From our analysis, we were able to verify which elements are the most frequent and relevant, and also what kind of information to expect when the users mention those elements. By resorting to that knowledge, we can craft interfaces that help the users to tell their stories in the most natural and effective way.

**Probable Transitions**
We define by analysing the stories, a idea of how the story is described by the user. Through the number of transitions that occur with great frequency and taking into account the element most likely to be the next to be remembered by the user, make the possibility to conclude which the next element of the story. This may help the natural course of the story, helping the memory of relevant information.

**INTERFACE**
In this section we will discuss aspects of the interface chosen, where users can attach narratives to their photos. Figure 5 represents the application interface after the user login. The user makes a pre-registration, which corresponds to completing a form with a series of data and then they can access their account. It is then presented the main screen of

**Figure 4 – Narrative system**

from the users in order to get the largest number of elements associated to the photo.

**Ambiguity**
Any ambiguity is common in stories. For instance, references to time, are often not very accurate with the user rather than referring specific dates, chooses a space of time. With this in account, the narrative will have some tolerance on the ambiguity. May be used techniques of disambiguation using contexts of other elements. The user refers, sometimes the element more than once in order to strengthen information and try to clarify it.

```
Place

The photo was taken in...
```

```
Let’s tell a story...

The photo was taken in Paris. Was taken One year ago. The device used was a Cell phone. The quality is Reasonable. The dimension is Average. The author of this photo was Me. No one was there. The photo can be described as notre dame catedral. The photo represents a Landscape. Was taken during trip to paris. The photo was taken in place.
```

```
End Story
```

```
Figure 5
```

```
```

```
```
We also measured the time users needed to retrieve photos. In terms of infrastructure, the site uses MySQL for store persistent information, using PHP and JavaScript programming language as well as HTML and CSS. It was also used AJAX to get server-side process without there being the need to do an update of the all page.

**EVALUATION**

In this section we present and analyze the results obtained after 40 people used the system. They were divided into two groups, in which one inserted tags in a traditional system of tagging, while the other group used the system based on stories presented and studied in this paper. The users entered new photos into the system in a three-month period.

The data was analyzed according to a set of metrics consisting on the amount of information that users inserted into the system, the time it takes to annotate each picture. We also measured the time users needed to retrieve photos.
We used this data to analyze the evolution of tags over time, in order to see whether narratives help increase tag reuse. Next we will present the main findings from the analysis of the two tagging systems and a comparison between them.

The set of photos that were used in the analysis of results consisted in 1229 photos. Of those, 52% were entered in the narrative-based version of the site, and 48% using the traditional tagging system. A 95% confidence t-test does not support this difference as being statistically significant.

Looking at the data from the narrative-tagging system, we see that users on average used 7.96 story elements in each story (stdev = 1.16). As each element is, in itself a tidbit of information describing the photo, we will consider it to be equivalent to a tag in the traditional system, for comparison purposes. However, it is to be noted that this does skew the comparison in benefit of traditional tagging, since several story elements contain more than one word, and if separated into individual tags, the actual number of those would be greater. We follow this conservative approach, though, as no deeper analysis of the story elements’ contents has been performed yet.

Users took an average of 122.45 seconds to tell a story (stdev = 18s). Users focused more on elements related to the Author, Device, Location, Quality and Size in their stories. One surprise was the frequency of the Time element in the stories, which was considerably lower than expected. Having been one of the most used elements in the interviews used to draw story patterns, the trend turned out not to affect the developed system.

In the system of traditional tagging, users have submitted an average of 1.35 tags per photo (stdev = 0.47), with an average of 45.6 seconds to insert tags in the photos (stdev = 10.79). It can be a high value of time to insert one or two tags, as the average may mean that users take great part of that time to think of words to associate to the content and that could help them recover the picture in the future. Apart from that, it is indeed a shorter time than that needed to tell a story, but the amount of information conveyed is much less as well. In essence, users are spending one third of the time to convey just one sixth of the information, even considering that an entire story element might correspond to a single tag.

In neither of the systems did we find statistically significant correlations between the time it took to annotate a photograph and the amount of information entered. This might be explained by the large time span the test took place in (3 months), that might have led to the users growing more accustomed with the site and improving their performance.

We see, thus, that by using narrative tagging users are able to convey much more information about their photos. That extra information might make the difference between being able to retrieve that photo at a later time or not.

In terms of tag reuse, taking into account the number of tags that were entered by users in both systems and the distinct tags in each system, we conclude that the narrative-based approach led to a greater reuse of tags. Using stories, the users, overall, entered 5363 elements in their stories. Of those, only 310 were distinct. This corresponds to 94% of story elements having been reused at least once.

On the other hand, for traditional tagging, the users entered a total of 793 tags, of which only 36% were used more than once. This shows that our approach is able to lead the users to far greater reuse of information than traditional approaches.

It might be argued that, since in the narrative-based approach some story elements provided the user with predetermined values to be chosen from, this would artificially increase tag reuse. That information is nevertheless relevant and was found to be spontaneous in the interviews. Removing those elements from the analysis does indeed lower the tag reuse percentage for narratives. Considering only free-form entry elements, we find a reuse of 89%, instead of the original 94%. Those elements did influence the results, but even accounting for that, narrative-based tagging still is, by far, able to elicit more consistent information from the users.

Something that was striking, from the viewpoint of the observer of the behavior of users, those who used the system of narrative felt less hesitant at the time of writing the tags, being guided through the dialogues that were emerging throughout the story.

Analyzing the data, characteristic problems of the traditional tagging like polysemy or synonyms appear like expected in the traditional system. In the narrative-tagging system those problems don’t appear as a result that tags are considered in a context, the element to which corresponds in the story. The tags presented in that context don’t show signs of ambiguity and there by problems like synonym and polysemy don’t appear.

**CONCLUSIONS**

This paper presents a new methodology that helps users in the process of tagging a photo. Instead of associating individual words to a photo, the user tells a story describing it. Telling stories allows the user to use a more structured approach, adding a set of contextual features to the description that help to retrieve the photo and that go beyond what can be achieved by traditional tagging. The story follows a structure, which was defined by analysing a set of interviews with users. That structure helps elicit more information for the user, resulting in richer descriptions.

Also, the narrative tagging method encompasses traditional tagging. Even if the users are at a loss regarding any other relevant information that stories would allow them to enter, they can still enter freeform text in fields such as “Description”, there placing what otherwise would be isolated tags. Most of the times, however, the semi-
structured nature of the storytelling process will lead the users to convey more information to the system.

We performed a user study in which 40 users, divided into two groups, evaluated our prototype. One used the system built with narratives and the other group a traditional tagging system.

In the results obtained it was possible to draw some conclusions concerning both systems. The user of the system based in narratives add 5.9 times more information than the traditional tagging systems user, which is a truly substantial increase of information. We found that although users are consuming in average more time to enter information, there wasn’t a proportional relationship observed between the amount of information added and the time taken to added it. In terms of tag reuse, it reached a value of 94% in the system based in narratives (89% after removing pre-determined elements from stories), against the 36% obtained in the traditional tagging system. Analyzing the tags quality we can see that the tags added in the narrative system are composed of several words and can help a lot when is time to retrieve the content. In the traditional tagging tags were preferably composed with one-word only, and in some cases it was used tags with two words together without space between them, which have little meaning in a community. Problems like polysemy and synonyms don’t appear in the tags added in the narrative-tagging system as a result that the tags are associated with a story element, which gives context to the tags removing possible ambiguity that might appear.

In short, we can conclude by stating that we have provided evidence supporting our initial hypothesis, that narrative-based semi-structured tagging of photos can, indeed, help the users to convey more and better information, alleviating some of the problems found in traditional tagging.

In terms of future work, we plan to perform a new user study where we will look closely at the photos entered by the users and study how particular tagging problems such as synonymy are affected by narrative tagging. We will also conduct tests using pre-defined sets of photos, to see to what extent can people describe photos that are not their own and about which they have little or no contextual information to relate to. This will lead to deeper insights regarding the use of narrative tagging in online communities and social networks.

Also, it has been pointed out that our approach could be adapted to children. Since their cognitive ability to associate a number of single words to content isn’t developed in their early stages of life, the structured process of tagging with narratives could facilitate thinking, facilitating the process of tagging for younger people.

ACKNOWLEDGMENTS

Tiago Guerreiro was supported by the Portuguese Foundation to a Science and a Technology bug. SFRH/BD/28110/2006.

REFERENCES


