# StoryTags: Once Upon a Time, There Was a Photo

**Nuno Tomás**
Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais, 49
1050-001 Lisboa, Portugal
Tel: +351-213100289
nmrt@ist.utl.pt

**Daniel Gonçalves**
Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais, 49
1050-001 Lisboa, Portugal
Tel: +351-213100289
daniel.goncalves@inesc-id.pt

**Tiago Guerreiro**
Instituto Superior Técnico
Universidade Técnica de Lisboa
Av. Rovisco Pais, 49
1050-001 Lisboa, Portugal
Tel: +351-213100289
tjvg@vimmi.inesc-id.pt


*CHI 2009*, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-246-7/09/04.

## Abstract
With the growing volume of digital information users must deal with, management and retrieval tasks have become increasingly problematic. A popular way to help users organize their information is *tagging*, as is the case in web sites such as flickr, delicious or youtube. Unlike traditional hierarchically-based organization principles, tagging is less strict and easier to employ. However, it is not without its own problems. Low tag reuse is just one of several issues that might hinder retrieval of a document or photo at a later time. We propose that narratives can provide a better way of tagging photos. Describing a photo by telling a story about it may yield more and better tags, as information in stories is organized as a structured, coherent whole. We present a prototype web application, *StoryTags*, that allows users to tell stories to tag their photos, and then to use those stories to retrieve them.

## Keywords
Tagging, narrative-based interfaces, digital photography

## ACM Classification Keywords
H.5.2 Interaction Styles, I.3.6 Methodology and Techniques: Interaction Techniques, H.3.3 Information

Search and Retrieval: Query formulation, H.5.m Miscellaneous.

## Introduction

With the advent of digital photography and video, coupled with broadband Internet connections, the number of digital documents and media that users must deal with on a daily basis is growing with hitherto unseen speed. All that information must be organized in order to facilitate its retrieval at a later time. However, traditional ways to organize digital items are hierarchy based and fraught with problems. In a hierarchy, all items must be classified into categories. However, each can be placed in a single one, not always being evident in which. Other times, several seem adequate. Furthermore, different applications require different, often incompatible, hierarchies. This does not properly reflect the way users recall their information, leading to retrieval problems.

A recent alternative to hierarchic classification schemes is *tagging*. Tagging is less structured than hierarchies [7]. The users describe their items resorting to a set of terms or keywords that they feel adequately describe them. Since several tags can be associated to a single document, it is possible for users to describe them in different ways, based on different meanings that an item can have, and anticipating different contexts in which it can be retrieved at a later time. As such, tagging is usually perceived as causing less cognitive load less than hierarchical classifications. Tagging is used, especially, in web applications such as Flickr and Delicious [5]. In those contexts, it is common to find tagging as a collaborative process, in which users can tag not only their items but also those of others. Often, this leads to the appearance of Folksonomies, sets of tags of preferential use by users of the community.

While attractive, tagging isn't, however, without its issues. Its strength, the lack of structure, is also its major problem. Three issues to take into account [1][4] are polysemy, the use of synonyms, and basic level variation. Polysemy occurs when a word has more than one sense. A "jaguar" can be a car or a feline. Hence, photos tagged as "jaguar" can be actually of two very different things. Likewise, the same concept can often be represented by more than one synonym. Finally, the same item can be described at more than one level. The photo of a lion can be tagged as "lion", "feline", or "animal". These problems lead to low tag reuse and might make the retrieval of items difficult, especially in collaborative contexts, in which we have little or no control over what others used as tags. Constraining the vocabulary that can be used when tagging would ultimately fail as it is impossible to predict all tags that might be required in all situations.

We propose that narrative-based tagging can help alleviate some of the aforementioned problems. If users tell short structured stories describing their items, this will help them recall more information, as stories are more than just sets of keywords but, instead, inter-relate the information items in them. This will help users recall more relevant information. Also, since they help structure the users' thoughts and memories, this might lead to more uniform descriptions of items and increase tag reuse.

In the following section we will describe a user study that allowed us to know what stories about digital photos, as told by the users, are like. Based on the

insights thus gained, we created StoryTags a collaborative web-site that uses narrative-based storytelling to let users annotate their digital photos.

## Stories about digital photographs

To help users tell meaningful stories about their digital photographs, it is important to know beforehand what to expect in those stories. This will allow the creation of an interface that allows users to tell them effortlessly.

We interviewed 20 users of different backgrounds. The only pre-requisite was that they had to regularly take digital photos, and use hosting sites such as Flickr or hi5 to store at least a subset of them. Two of the users were professional photographers. Each user was asked to describe, in their own words, three of his/her photographs. Thus, overall, 60 stories about digital photos were collected. After ensuring that the users' privacy would be taken into account, all interviews were recorded for future analysis.

The interviewer did not intervene in the storytelling process, unless the users seemed at a loss as what to mention next. In that case, a set of pre-determined phrases was used to help the users move forward. We took care to note when this happened, and the information mentioned right after the prompting by the interviewer was considered to have been induced by the interaction, while the rest was considered to be spontaneous [2][3].

The stories were transcribed and subjected to contents analysis. The different phrases were classified into several possible elements, describing different characteristics of photos. This was done by hand, as it would impossible to predict all possible information in stories and automate its classification.

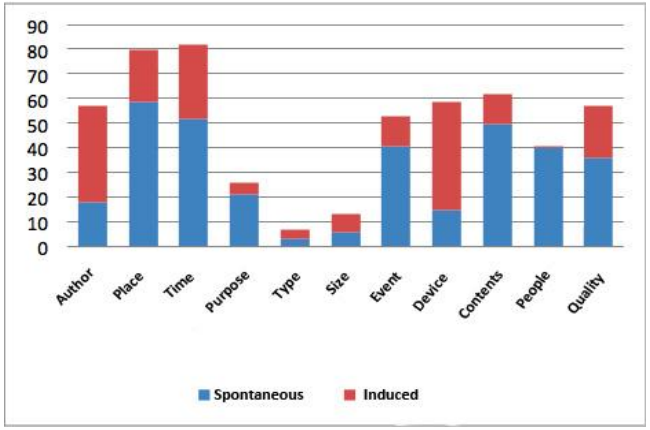


**figure 1.** number of story elements found

We found that stories about digital photos are composed of the following elements, in decreasing order of frequency: Time (when the photo was taken); Place (where); Contents (a short description of what is depicted in the photo); Device (what camera/phone it was shot with); Quality (whether it is a good, interesting photo or not); Author (who took the photo); Event (the event at which the photo was taken); People appearing in the photo; Purpose (what the photo is to be used for); Size; and Type (landscape or portrait). As can be seen in **figure 1**, Time and Place were by far the most frequent, with an overall frequency of around 80. As there were only 60 stories in the study, this means they are so important to users that they are mentioned more than once in stories. On the other hand, Purpose, Type and Size were seldom mentioned. It is also interesting to see that Author and Device occur, most times, as induced elements, which means

that the users recall this information with ease but only after being prompted to do so.

This suggests that *establishing dialogues* with users is important to elicit relevant information from them. Also, we found little evidence for the *need for personalization*, in terms of gender, age and other such factors, as stories told by users of different types were similar (although an extended study might be necessary to prove this). Also, there is some level of *ambiguity* in stories, often when mentioning Time, referring to holidays, seasons, etc., showing the need for the use of commons sense knowledge when interpreting stories.

Finally, we also performed a relational analysis, where the order in which the different story elements succeeded each other in the stories was analyzed. This allows us to determine the most likely element transitions and train a hidden-markov model [6] to create an archetypical story structure: *Place, Time, Time, Device, Quality, Size, Author, People, Contents, Purpose, Type and Event.* This was the basis for the StoryTags prototype described in the following section.

## The StoryTags Interface

The interface of the StoryTags application is depicted in **figure 2**. The web site is organized as most common Web 2.0 sites. Each user has its own personal area that can be used to upload and manage their digital photos.

A user's home page shows the photos that have been uploaded. Those that still haven't been annotated by the user are highlighted with a warning message. This provides an immediate visual clue as to what still needs to be tagged. It is also possible to browse all photos

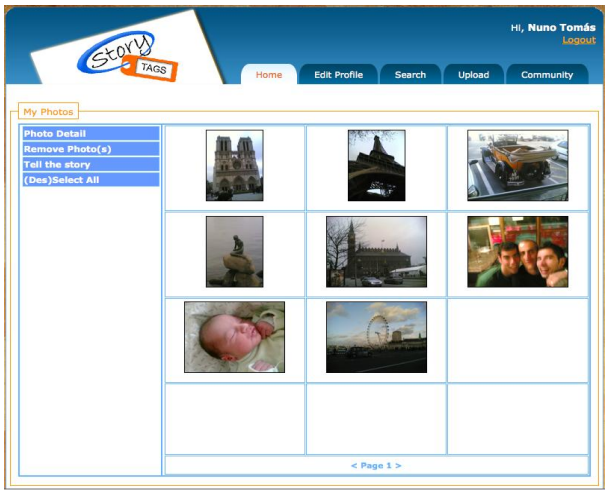


**figure 2.** StoryTags photo collection view

uploaded in the system, and perform the usual tasks of changing their description, deleting them, etc.

The user can upload one or several photos at once. Whenever this is done, the user will be prompted to tell a story describing one or more of those photos. Each photo can be described by more than one story. This allows users to tell stories that describe entire sets of photos ("these are my holiday in Greece photos"), and then more specific stories describing smaller groups or individual photos. This makes performing at least a simple tagging of them more efficient.

When telling a story, the screen is divided into three main areas (**figure 3**). The largest, to the right, is the story area, in which the story is created. Below it, the photos being described are shown. To the left, several specialized dialogue boxes are presented in succession, allowing the users to add new information to the story.


**figure 3.** StoryTags storytelling interface

The storytelling process works in the following way: for each story element, the users are presented with an incomplete sentence that lets them know what element should be mentioned next. At the same time, a dialogue box, on the left, will provide the means for the entry of the relevant information for that element. The dialogues were directly derived from the stories collected in the user study described in the previous section. They are sufficiently expressive to allow all the information in those stories to be entered by users. When a user enters new information using the dialogue, the incomplete sentence changes, reflecting the information just entered, and the next sentence, corresponding to the next story element is suggested to the user. A sample story told using this method can be found in **figure 4**.

The order in which the different story elements are presented to the users is the one found to be most

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 .

**figure 4.** sample StoryTags story

likely from the analysis of the interviews. However, the users have the liberty to change it as they see fit, and the interface will learn and adapt to that users' particular storytelling process.

StoryTags also allows narratives to be used as a retrieval mechanism. Besides the usual keyword-based search, the users can tell stories describing photos they want to find, and those stories will be compared to those annotating the photos, in search of matches.

We chose this way to allow users to tell their stories because it was found, in the interviews, that dialoguing with users is important to drive the storytelling process forward. By suggesting different story elements to the users we accomplish this. Also, since it is impossible to predict exactly what information will be in the stories, allowing users to tell them using free-form text would lead to hard to solve natural language understanding problems. Furthermore, in the interviews it became clear that few users would have the patience to completely write down the entire story. Our fill-in-the-blanks approach addresses this issue by requiring users to enter only relevant information about photos.

## Evaluating Narrative-Based Tagging

In order to validate that our approach an extensive user study is now ongoing. The prototype has a mode that allows tagging to be performed mimicking the

method used in Flickr. Twenty users are tagging their photos using narratives, and twenty other are doing it using the Flickr approach. Over the course of two months, they are being asked to weekly upload at least 20 of their photos and tag them. Also, every two weeks they must tag a set of 20 photos chosen by us, representing well know public places, events and people. This will allow us to measure tag reuse for each user and across users, see if there are any differences between tags for personal photos and others, and directly compare both tagging methods.

## Conclusions

Tagging has been increasingly used as an alternative to the hierarchical organization of digital items. Its flexibility makes annotating those items easier. However, it is not without problems. The use of synonyms, basic level variation and polysemy can lead to low tag reuse and retrieval problems. We propose that telling stories to describe items in general and digital photos in particular might help alleviate those problems, while maintaining the advantages that make tagging so appealing. User studies allowed us to get a thorough characterization of stories describing digital photos. Based on those results, we created a prototype system, StoryTags, that allows users to do that in a collaborative setting. StoryTags is currently undergoing an extended user study in which we will be able to compare traditional and narrative-based tagging, showing the latter's advantages.

## References

[1] Golder, S., and Huberman, B. A.The Structure of Collaborative Tagging Systems. HP Labs technical report, 2005. Available at http://www.hpl.hp.com/research/idl/papers/tags/

[2] Gonçalves, D., Jorge, J. Evaluating Stories in Narrative-Based Interfaces. In Proc. Int. Conference on Intelligent User Interfaces 2006, ACM Press, 29 Jan - 1 Feb 2006, Sydney, Australia.

[3] Gonçalves, D., Jorge, J. "Tell me a Story": Issues on the Design of Document Retrieval Systems. In Proc. DSV-IS'04, Lecture Notes on Computer Science, Springer-Verlag, July 2004, Hamburg, Germany.

[4] Mathes, A. Folksonomies – Cooperative Classification and Communication Through Shared Metadata. UIC Technical Report, 2004.

[5] Marlow, C., Naaman, M., Boyd, D., Davis, M., HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read. August 2006 HYPERTEXT '06: Proceedings of the seventeenth conference on Hypertext and hypermedia

[6] Rabiner, L. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, vol 77, no 2, February 1989.

[7] Hayman, S., Folksonomies and Tagging: New developments in social bookmarking. Ark Group Conference: Developing and Improving Classification Schemes, 27-29 June, 2007, Rydges World Square, Sydney, Australia.