User-centric Metadata for Mobile Photos

Risto Sarvas
Helsinki Institute for Information Technology HIIT
P.O.Box 9800
02015 TKK, Finland

risto.sarvas@hiit.fi

ABSTRACT
Consumers are overwhelmed by the amount of personal media they create themselves and the media others share with them. They are gradually facing the problem how to use, manage, and organize all this personal media. Public and commercial media archives have traditionally addressed this media management problem by having metadata describing the content of the media. This approach, however, has two inherent problems in the domain of consumer media: lack of end user motivation for annotation, and the dynamic and semantic nature of personal content metadata ontologies. In this paper we describe our empirical and user-centric approach into this issue. Based on our studies with mobile photo systems, we argue that consumer content metadata should be a side product of some other user activity like photo sharing or social discourse. To design user-centric metadata ontologies we need to understand both the uses people have for the media and the system design and implementation decisions that affect how and for what purposes the media is created.

Categories and Subject Descriptors
H.5.1 [**Information interfaces and presentation** (*e.g.*, HCI)]: Multimedia; H.4.3 [**Information systems applications**]: Communications Applications.

General Terms
Design, Human Factors.

Keywords
Camera phones, media content metadata, personal media.

1. INTRODUCTION
The availability of media creation devices such as digital cameras and digital video cameras combined with broadband Internet access has made the creation and sharing of personal media an everyday activity. The result is that people are being gradually overwhelmed by the amount of media they create themselves and shared by others, and they are facing the problem of media management: how to organize, handle, browse, and access the gigabytes of personal digital pictures and videos.

By *personal media* we mean media created mainly for personal use, not for commercial purposes. Of all the personal media we focus on photographs which have been and still are the most popular media people create themselves (excluding perhaps text in emails).

In commercial and public media archives, such as libraries and media company archives, the issue of media management has been addressed by annotating the pictures with metadata describing the contents (*e.g.*, IPTC News Markup Language). Therefore, the background for addressing the personal media management problem is in the domain of content describing metadata for public and commercial archives.

A typical metadata ontology for personal image management consists of two levels: People, Events, and Places as the top-level categories, and the individual metadata, that is, the people, events, and places in the pictures (*e.g.*, Adobe Photoshop Album). For example, one photograph may have the metadata:

> **People**: Uncle Matt, Roger, Mike, John Smith.
> **Event**: Mike's 13th birthday.
> **Places**: Home.

A more sophisticated ontology might have several layers such as a hierarchy of places: Country – City – Street.

However, using personal metadata, such as above, to address people's personal media management problem has two inherent issues that do not manifest themselves as such in the case of public or commercial archives:

- End user **motivation** for entering the metadata.
- **Dynamic and semantic nature** of personal metadata.

In this paper we describe these two problems inherent in user-created and managed content metadata, and present our empirical and user-centric approach in researching these issues. Our focus is on mobile phone photography, because camera phones provide an exceptionally promising platform for the creation of metadata. In the end we conclude that to facilitate the creation of content describing metadata, it should be created as a *side product* of some other end user activity, and that to design metadata ontologies for consumer use there has to be a good understanding of the *uses* people have for personal media and the how the system implementation affects this use.

2. USER-CREATED METADATA
Unlike in public or commercial archives, personal media metadata is often created, managed, and used by an individual or a small group such as a family. The metadata is most often an organizing feature in a digital photograph management system such as Adobe Photoshop Album or Apple iPhoto. These systems provide a basic ontology for the user to start the metadata annotation (see the example above), and also a possibility to extend the ontology further (*e.g.*, to add a category like *Cars* if the user is a car enthusiast, or to sub-divide an existing category like *Events* into *Family Events*, *Work Events*, etc.).

Both the metadata structure (*i.e.*, the ontology) and the metadata itself are personal and semantic information that can not be easily

acquired by mere computation. Computational extraction and generation of information is often limited to information that has little value in personal image management (*e.g.*, camera exposure time, or image color histogram). This semantic gap between the information acquired automatically and the information valuable to the user means that creating personal content metadata requires user participation and can not be fully automated.

2.1 End User Motivation

In the public media archive domain, and even more in the domain of commercial media archives, the cost of annotating media with content metadata has to be less than the benefits. In the case of public or commercial archives the annotation work is often part of the job and, therefore, motivated to the end user in that sense.

In annotating personal media it is also a case of costs and benefits. However, in personal media annotation there is no one paying to do the time-consuming work. Therefore, the end users motivation is based on the direct personal benefits of having descriptive metadata. The benefits are by no means clear. First of all, for a consumer the concept of metadata is not trivial, especially when presented as a structure with relationships and hierarchies. Secondly, people do not necessarily understand the need for annotations. One interviewed user commented on her paper photo albums: "Why should I write down who is in the picture? I remember everyone, and if I forget who someone is they are probably not that important anyway."

The problem is that the benefits of metadata are not immediate: the annotation task is often separate from the tasks that would demonstrate the benefits of metadata (*e.g.*, searching for pictures or applications taking advantage of metadata). Also, people's paper photographs are often notoriously disorganized, so motivating them to organize their digital photos is not on solid ground.

2.2 Dynamic Nature of Personal Metadata

The other main issue with user-created metadata is related to its personal and semantic nature. These characteristics are inherent not only in the metadata information but also in the metadata structure. Depending on the individual doing the annotation the metadata information can be very different. For example, a person in a picture can be annotated as "John Smith", "John", "brother", or "John from work". Also, the ontologies can differ significantly, like the granularity and structure of locations. For example, for a tourist "Japan" might suffice as a location but for a resident of Tokyo the locations are more detailed and structured.

Secondly, personal media metadata reflects the user's personal life: people, places, events etc. This information is dynamic by nature because relationships to people and places change over time (*e.g.*, the tourist who visited Tokyo moves to live in Tokyo and this probably changes the past, present, and future picture metadata and ontology). Thirdly, different people call the same things by different names, and use the same name for different things. This vocabulary problem is true even when people use the same language.

These problems mean that personal metadata is not standard and not easily shared, and that the task of managing the dynamic metadata and ontologies is another burden to the end user.

3. USER-CENTRIC APPROACH

Current personal content metadata ontologies, like the examples above, have their background in the media ontologies familiar from public and commercial archives. In these archives the media management systems are *metadata-centric*. In consumer media management the user gives little value to the metadata as such, therefore consumer media metadata should be more *user-centric*.

In our research we have approached the design of user-centric content metadata by a combination of qualitative user studies on people's photography habits and construction of prototype systems that are then evaluated by user interviews. As a technological platform we have focused on mobile phones with cameras. With their inherent network connectivity, access to personal and contextual information, open programming interfaces, and increasing popularity mobile phones with cameras are a promising device for personal media creation, sharing, and metadata annotation.

3.1 The Mobile Media Metadata System

Our first system designed for user evaluation, named *MMM-1* for Mobile Media Metadata version 1, focused on facilitating the picture metadata annotation process [1]. Based on the available contextual metadata on the mobile phone and similarity in previously inserted metadata on a remote server, the system suggested content metadata to the user immediately after capture (see Figure 1). This had the benefit of facilitating the time consuming task of annotation by generating content metadata semi-automatically. The *MMM-1* system was evaluated by user interface testing, surveys, and interviews made with over 40 researchers and students at the University of California at Berkeley, USA [4].

The test results supported the issues discussed above. Firstly, without a direct application for the annotated metadata the users had a hard time understanding the reasons behind the annotation task. Secondly, the annotation effort had to be minimal – even the smallest of delays or errors in the user interaction quickly became frustrating. Thirdly, the users' metadata showed the inconsistency of annotations between people due to the personal and semantic differences in describing the contents.

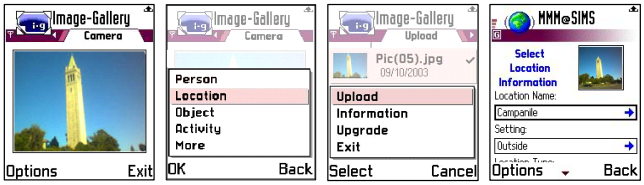

Figure 1. The phone user interface of *MMM-1*. Immediately after capturing the image the user is asked what the picture is about, then the image and available metadata is uploaded to a server, and the server returns its guesses for verification.

3.2 The MobShare System

The second system, *MobShare*, was designed based on user interviews and literature on camera phone use and snapshot photography [3]. The goal of the system was to gain more understanding in what purposes the camera phones are used for and what is their relation to people's photography habits and social networks. The main characteristics of *MobShare* were that the pictures were shared in galleries on a web page (*i.e.*, the captured photos were uploaded from the phone to a web server), and each gallery had a list of recipients chosen by the user from the phone's address book (see Figure 2). On the web page part of the system, users had individual accounts, and their own and shared pictures were organized by galleries, time of capture, and owner (see Figure 3).

The *MobShare* system was evaluated by a six week user test with 52 participants [2]. The five primary users were a group of friends who shared 525 of 589 of their captured mobile photos with 48 people consisting of their friends, relatives, and colleagues. We observed the kind of pictures taken (70 % had a person, 21 % an object, 13 % a scenery, and 12 % a dog), the galleries created (53 % were about events, 16 % themes, 11 % travels, 9 % exploring the system, and 11 % were "other"), and the social discourse around the shared pictures (in-group post-event discourse, self-documents and reports, greetings and thanks). We also noticed that the selection of recipients for galleries was surprisingly dynamic depending on the contents of the pictures and the communications the user had had with the recipient prior to sharing.

3.3 Heterogeneity of Mobile Imaging Systems

Camera phones are often viewed simplistically as digital cameras with poor picture quality. This view neglects the wide variety of uses people have for mobile photos which can be seen in the wide variety of systems for mobile imaging. Unlike traditional film or digital photos, mobile pictures are often a part of a networked computer system: Mobile photos are used for general messaging and publishing (multimedia messaging, instant messaging, and photo blogging), or are the result of an end-to-end applications such as *MMM-1* or *MobShare*. This means that technical implementations of each system strongly affects what kind of pictures people take, to whom they share them, and what kind of social discourse emerges around them [2]. This heterogeneity of systems means that a metadata ontology designed based on one system's use can not be generalized without identifying the implementation effects of that particular system. Therefore, the design of content metadata for mobile media must take into account, not only the new uses people have for mobile media, but also the systems used.

Figure 2. The sharing of pictures in the *MobShare* phone client. First the pictures are selected for posting, and then a new gallery/album is created for the selected pictures. The gallery is named and the recipients are selected. In the end the pictures are optimized and uploaded to the web server.

4. CONCLUSIONS

We have identified two problems in media content metadata for personal use: end user motivation for annotation, and the dynamic and semantic nature of personal content metadata ontologies. For these reasons rigorous ontologies and explicit annotation tasks familiar in public and commercial media management do not work in consumer applications. Based on our studies with mobile photo systems, we argue that to minimize the user's annotation burden, the content metadata should be a side product of some other user activity like photo sharing or social discourse. To design user-centric metadata ontologies that are annotated and managed as a side product, we need to understand both the uses people have for the media, and the system implementation decisions that affect how and for what purposes the media is created.

Our future work is designing a personal media metadata ontology for adressing the issues discussed here. The ontology will support the two most common uses for personal media: organization and sharing. To address the issues dicussed above, the ontology will have two layers: a lower layer for describing standard information such as location coordinates, capture and sharing time and date. An upper layer would be reserved for more semantic and personal information such as names of people, events, and location, as well as, groups of recipients for shared media and so on.

5. REFERENCES

[1] Sarvas, R., Herrarte, E., Wilhelm, A., Davis, M. Metadata Creation System for Mobile Images, In *Proc. of MobiSys 2004.* ACM Press, New York, NY, 2004, 36-48.

[2] Sarvas, R., Oulasvirta, A., Jacucci, G. Building Social Discourse Around Mobile Photos – A Systemic Perspective. In *Proc. Of MobileHCI 2005* (Forthcoming 2005).

[3] Sarvas, R., Viikari, M., Pesonen, J., Nevanlinna, H. MobShare: Controlled and Immediate Sharing of Mobile Images, In *Proc. of Multimedia 2004*. ACM Press, New York, NY, 2004, 724-731.

[4] Wilhelm, A., Takhteyev, Y., Sarvas, R., Van House, N., and Davis, M. Photo Annotation on a Camera Phone. In *Proc. of CHI2004*, ACM Press, 2004, 1403-1406

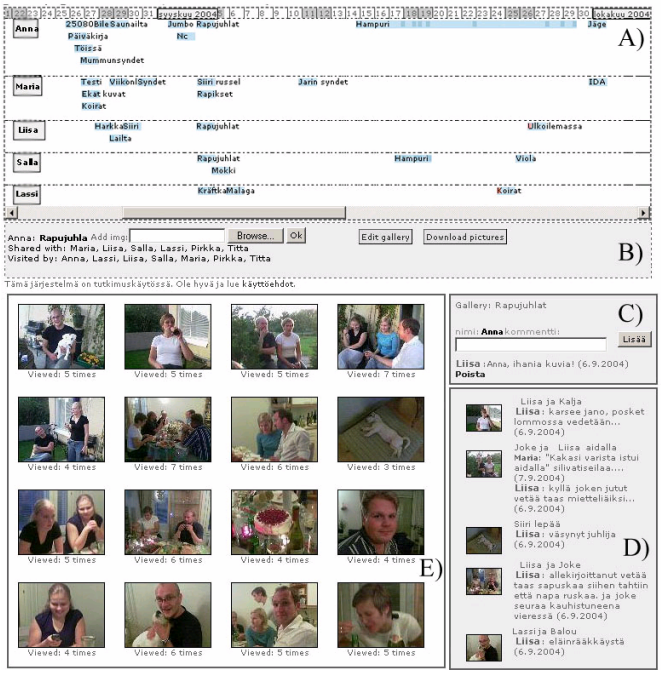

Figure 3. The web page interface of *MobShare*. A) User's own galleries, and galleries shared with her. B) List of people who the gallery is shared with and who have visited it. C) Gallery-level comments. D) Picture-level comments. E) The thumbnails of pictures in the gallery.