RISTO SARVAS

DESIGNING USER-CENTRIC METADATA FOR DIGITAL SNAPSHOT PHOTOGRAPHY

Helsinki University of Technology
Department of Computer Science and Engineering
Software Business and Research Institute

and

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

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Computer Science and Engineering for public examination and debate in Auditorium T2 at Helsinki University of Technology (Espoo, Finland) on the 15th of December, 2006 at 12 o'clock noon.

HELSINKI UNIVERSITY OF TECHNOLOGY P.O. BOX 1000, FIN-02015 TKK http://www.tkk.fi	ABSTRACT OF DOCTORAL DISSERTATION
Author: Risto Sarvas	
Name of the dissertation: Designing User-Centric Metadata for Digital Snapshot Photography	
Date of manuscript: 12.6.2006	Date of the dissertation: 15.12.2006
Monograph ☐	Article dissertation (summary + original articles) ☒
Department: Computer Science and Engineering Laboratory: Software Business and Research Institute Field of research: Digital Media Opponent(s): Professor Walter Bender Supervisor: Professor Reijo Sulonen Instructor: Professor Marko Turpeinen	

Abstract

With digital cameras, camera phones, and camcorders connected to personal computers and the Internet, people are capturing, sharing, and storing more and more media created by themselves. This has created the organizing challenge called media management problem: how to browse and find media from the growing personal media collections. Metadata is seen as one of the most promising technologies to address this problem. Previous work on metadata design has focused on public or commercially produced media - not on media created by consumers for themselves or their social networks.

Taking snapshot photography as a starting point, this dissertation looks into the media management problem from a user-centric design perspective. In this work, three consecutive systems were designed: *MMM-1*, *MobShare*, and *PhotosToFriends*. Each of these systems was evaluated in extensive user trials. The objective of the user trials was to provide a better understanding of the uses people have for mobile photos and metadata. The results gained from constructing the systems and the associated user trials are divided into three. First, they contribute to understanding social uses for mobile photographs: what was photographed, with whom were the photos shared, and what kind of social activity emerged. Second, the results can be applied into designing systems for photo sharing: the use of galleries, control over sharing, support for discussions, and social awareness and notifications. Third, what are the implications for metadata that the social uses and system design have.

The main implications are the inherent problem of personal photo information being contextual, dynamic, and highly semantic, and the strong coupling of metadata and its application. To address these problems we propose the concept of social metadata, which takes advantage of the social activity in photo sharing systems. The user-centric design approach also brings forth that content metadata generation should be approached from several angles - social metadata, user tagging and contextual information - in addition to the more traditional content-based analysis.

Keywords: mobile photography, digital photography, metadata, personal media, snapshot photography, photo sharing	
UDC	Number of pages: 130 + 51 (original articles)
ISBN-13 (printed) 978-951-22-8443-6 ISBN-10 (printed) 951-22-8443-X	ISBN-13 (pdf) 978-951-22-8444-3 ISBN-10 (pdf) 951-22-8444-8
ISSN (printed) 1458-9451	ISSN (pdf) 1458-946X
Publisher: Helsinki Institute for Information Technology HIIT, Helsinki University of Technology TKK	
Print distribution	
The dissertation can be read at http://lib.hut.fi/Diss/	

TEKNILLINEN KORKEAKOULU PL 1000, 02015 TKK http://www.tkk.fi	VÄITÖSKIRJAN TIIVISTELMÄ
Tekijä: Risto Sarvas	
Väitöskirjan nimi: Käyttäjäkeskeisen metatiedon suunnittelu digitaaliseen näppäilykuvaukseen	
Käsikirjoituksen jättämispäivämäärä: 12.6.2006	Väitöstilaisuuden ajankohta: 15.12.2006
Monografia ☐	Yhdistelmäväitöskirja (yhteenveto + erillisartikkelit) ☒
Osasto: Tietotekniikan osasto Laboratorio: Ohjelmistoliiketoiminnan ja -tuotannon laboratorio Tutkimusala: Digitaalinen media Vastaväittäjä: Professori Walter Bender Työn valvoja: Professori Reijo Sulonen Työn ohjaaja: Professori Marko Turpeinen	

Tiivistelmä

Digitaalikameroiden, kamerapuhelimien ja digitaalisten videokameroiden ollessa nykyään yhdistettynä kotitietokoneiseen ja sitä kautta Internetiin ihmiset luovat, jakavat ja tallentavat entistä enemmän itse tekemäänsä mediaa. Tämän seurauksena on syntynyt organisointihaaste, jota kutsutaan henkilökohtaisen median hallintaongelmaksi: kuinka selata mediaa tai löytää sieltä haluamansa media? Metatietoa pidetään yhtenä lupaavimmista teknologisista ratkaisuista tähän ongelmaan. Aikaisempi tutkimus metatiedon suunnittelussa on keskittynyt julkiseen tai kaupalliseen mediaan – ei mediaan, jonka kuluttajat luovat itselleen tai sosiaaliselle verkostolleen.

Tämä väitöskirja tutkii median hallintaongelmaa käyttäjäkeskeisestä näkökulmasta, jossa lähtökohtana on näppäilyvalokuvaus. Tässä työssä rakennettiin kolme peräkkäistä järjestelmää: *MMM-1*, *MobShare* ja *PhotosToFriends*. Kukin näistä järjestelmistä evaluoitiin kattavilla käyttäjäkokeilla, joiden tavoitteena oli laajentaa ymmärrystä ihmisten mobiilikuvien käytöstä sekä metatiedosta. Järjestelmien rakentamisesta ja niihin liittyvistä käyttäjäkokeista saadut tulokset voidaan jakaa kolmeen osaan. Ensiksi, tulosten kautta voidaan paremmin ymmärtää mobiilikuvien sosiaalista käyttöä: mitä kuvattiin, keille kuvia jaettiin ja minkälaista sosiaalista toimintaa syntyi. Toiseksi, tuloksia voidaan käyttää kuvienjakojärjestelmien suunnittelussa: gallerioiden käyttö, jakamisen kontrollointi, keskustelun tukeminen sekä tietoisuus ja tiedottaminen sosiaalisesta aktiivisuudesta. Kolmanneksi, mitä seuraamuksia kahdella edellisellä kohdalla on metatiedon suunnittelulle.

Tärkeimmät seuraamukset ovat se, että henkilökohtainen tieto on kontekstuaalista, dynaamista ja semanttista, ja metatiedon sekä sitä käyttävien sovellusten välillä on vahva keskinäinen liitos. Näiden ongelmien ratkaisemiseksi ehdotamme sosiaalisen metatiedon käsitettä, joka käyttää hyväksi kuvienjakojärjestelmien sosiaalista aktiivisuutta. Käyttäjäkeskeinen lähestyminen suunnittelussa nostaa esille myös, että sisältöä kuvaavan metatiedon luomista pitäisi lähestyä useasta eri näkökulmasta – sosiaalinen metatieto, käyttäjien ”tägit” ja konteksti-informaatio – perinteisen sisältöpohjaisen analyysin lisäksi.

Asiasanat: mobiilivalokuvaus, digitaalivalokuvaus, metatieto, henkilökohtainen media, näppäilykuvaus	
UDC	Sivumäärä: 130 + 51 (erillisartikkelit)
ISBN-13 (painettu) 978-951-22-8443-6 ISBN-10 (painettu) 951-22-8443-X	ISBN-13 (pdf) 978-951-22-8444-3 ISBN-10 (pdf) 951-22-8444-8
ISSN (painettu) 1458-9451	ISSN (pdf) 1458-946X
Julkaisija: Tietotekniikan Tutkimuslaitos HIIT, Teknillinen Korkeakoulu	
Painetun väitöskirjan jakelu	
Luettavissa verkosta osoitteessa http://lib.hut.fi/Diss/	

ORIGINAL ARTICLES

This dissertation consists of a summary and five publications. The publications, presented below with abstracts and descriptions of the author's contribution, are referred to in the text by their Roman numerals I-V.

I. Sarvas, R., Herrarte, E., Wilhelm, A., and Davis, M. 2004. Metadata creation system for mobile images. In *Proceedings of the 2nd international Conference on Mobile Systems, Applications, and Services* (Boston, MA, USA, June 06 - 09, 2004). MobiSys '04. ACM Press, New York, NY, 36-48.

Abstract. The amount of personal digital media is increasing, and managing it has become a pressing problem. Effective management of media content is not possible without content-related metadata. In this paper we describe a content metadata creation process for images taken with a mobile phone. The design goals were to automate the creation of image content metadata by leveraging automatically available contextual metadata on the mobile phone, to use similarity processing algorithms for reusing shared metadata and images on a remote server, and to interact with the mobile phone user during image capture to confirm and augment the system supplied metadata. We built a prototype system to evaluate the designed metadata creation process. The main findings were that the creation process could be implemented with current technology and it facilitated the creation of semantic metadata at the time of image capture.

As the first author, Risto Sarvas has, together with co-authors, designed the presented approach for metadata creation, designed and managed the implementation of the associated MMM-1 system, and contributed to the design of

the user tests. He has written the whole article with suggestions and comments from the co-authors.

II. Wilhelm, A., Takhteyev, Y., Sarvas, R., Van House, N., and Davis, M. 2004. Photo annotation on a camera phone. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM Press, New York, NY, 1403-1406.

Abstract. In this paper we describe a system that allows users to annotate digital photos at the time of capture. The system uses camera phones with a lightweight client application and a server to store the images and metadata and assists the user in annotation on the camera phone by providing guesses about the content of the photos. By conducting user interface testing, surveys, and focus groups we were able to evaluate the usability of this system and motivations that will inform our development of future mobile media annotation applications. In this paper we present usability issues encountered in using a camera phone as an image annotation device immediately after image capture and users' responses to use of such a system.

Risto Sarvas's contribution to this article is, together with co-authors, the design and implementation of the user tests. He has contributed significantly to writing the introduction, methodology, system overview, and user interaction challenges.

III. Sarvas, R., Viikari, M., Pesonen, J., and Nevanlinna, H. 2004. MobShare: controlled and immediate sharing of mobile images. In *Proceedings of the 12th Annual ACM international Conference on Multimedia* (New York, NY, USA, October 10 - 16, 2004). MULTIMEDIA '04. ACM Press, New York, NY, 724-731.

Abstract. In this paper we describe the design and implementation of a mobile phone picture sharing system MobShare that enables immediate, controlled, and organized sharing of mobile pictures, and the browsing, combining, and discussion of the shared pictures. The design combines research on photography, personal image management, mobile phone camera use, mobile picture publishing, and an interview study we conducted on mobile phone camera users. The system is based

on a client-server architecture and uses current mobile phone and web technology. The implementation describes novel solutions in immediate sharing of mobile images to an organized web album, and in providing full control over with whom the images are shared. Also, we describe new ways of promoting discussion in sharing images and enabling the combination and comparison of personal and shared pictures. The system proves that the designed solutions can be implemented with current technology and provides novel approaches to general issues in sharing digital images.

Risto Sarvas has, together with co-authors, designed the presented MobShare system. He has been the main contributor to the analysis of the system, and he has written the whole article with support from the co-authors.

IV. Sarvas, R., Oulasvirta, A., and Jacucci, G. 2005. Building social discourse around mobile photos: a systemic perspective. In *Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Services* (Salzburg, Austria, September 19 - 22, 2005). MobileHCI '05, vol. 111. ACM Press, New York, NY, 31-38.

Abstract. Camera phones have been viewed simplistically as digital cameras with poor picture quality while neglecting the utility of the two key functionalities of mobile phones: network connection and access to personal information. This is the first HCI paper to ex-amine mobile photos from a systemic perspective: how assignment of phases of mobile photo lifecycle to different platforms affects social discourse around shared photos. We conducted a 6-week user trial of MobShare, a tripartite system with dedicated functions and task couplings for a mobile phone, a server, and a PC browser. We analyze how MobShare's couplings and distribution of functionalities affected the observed types of social discourse that formed around mobile photos: in-group post-event discourse, self-documents and reports, greetings and thanks. Several central design issues arising from the systemic view are discussed: heterogeneity of environments, integration and distribution of functionalities, couplings and decouplings of interaction tasks, notification mechanisms, and provision of necessary UI resources for different tasks.

As the first author, Risto Sarvas has, together with co-authors, formed the analysis approach of looking at mobile photos from a systemic perspective. He has gathered all the data, and with co-authors, analyzed the data. He has significantly participated in the writing of the article.

V. Sarvas, R., and Turpeinen, M. Social Metadata for Consumer Photography. Submitted to *IEEE Computer* on October 15th, 2006.

Abstract. The increasing number of media creation devices is adding to the organization and media management problem for personal media. Content describing metadata is seen as the most prominent answer to this problem. However, current metadata standards have shortcomings when it comes to the domain of personal media. This is because the standards are designed for public and professional uses of media and in personal media some fundamental principles about public and professional media and metadata do not hold true. Based on our studies and literature on people's photography habits we propose to use social information as a metadata resource for storing, retrieving, and enabling new applications. We argue that information on social activity around personal media is both important to the end-users and enables new kinds of applications for solving the personal media management problem, and that future work on personal media metadata should take into account the social information related to the media.

Risto Sarvas's contribution to the article, together with the co-author, is the describing the problem setting, the system involved, and creating the concept of social metadata. As the first author, he has significantly participated in the writing of the article.

Table of Contents

FOREWORD

I began my doctoral studies in December 2001 part of the Tekes project *Digital Economy Core* (DECore) at the Helsinki Institute for Information Technology HIIT. In August 2002 I continued my work as a visiting scholar at the School of Information Management and Systems (SIMS) at the University of California, Berkeley. During the year 2003 my work began in the Tekes project *Mobile Content Communities* (MC2) at HIIT and the collaboration with SIMS was part of it. I returned from the U.S. in December 2003 and during 2004 got involved also in the Tekes project *Rich Semantic Media for Private and Professional Users* (RISE). Since August 2005 my work has been in the KCL and Tekes future snapshot photography project *Immortalidad*. During these years my doctoral work has been also funded by the Graduate School for Electronic Business and Software Industry (GEBSI) and by the Kaupallisten ja Teknillisten Tieteiden Tukisäätiö (KAUTE) both of which I thank very much for their support. I am also very grateful to all the companies who were partners in the above-mentioned projects (Accenture, Alma Media, Asianajajien Teknologiaoikeudellinen Tutkimusyhdistys, Codetoys, Elisa, LM Ericsson, Futurice, KCL, Nixu, Nokia, Profium, Starcut, Sulake, Sumea, Suomen Posti, Swelcom, TeliaSonera, Veikkaus, Yliopistopaino, and Yleisradio). Special thanks naturally go to the Finnish Funding Agency for Technology and Innovation (Tekes) for making this kind of research possible in the first place.

My work would not have been possible without the help, encouragement, enjoyable company, and wisdom of several people. First of all I would like to acknowledge the guidance and support from my instructor Professor Marko Turpeinen. Not only has he been the main guide and support for me through this process but he is also responsible for creating an exceptionally friendly and inspiring

working environment in the Digital Content Communities (DCC) research group at HIIT. On the same note, I would like to thank Professor Martti Mäntylä for always having time to listen to a variation of problems or ideas, and creating and directing a research institute that is the best place possible to do research on information technology. I am also grateful to my supervisor Professor Reijo Sulonen for helping out in the last stages of wrapping-up and finalizing this thesis. Professor Jukka Kemppinen I thank for teaching me whole new perspectives to studying technology and for excellent references to military technology as a source for understanding technology in history. This work would not have happened without the extraordinary inspiration and drive of Professor Marc Davis, who introduced me to multimedia and metadata. The teachings and discussion with Professor Nancy Van House have also been essential during all these years. Special thanks go also to the pre-examiners of this thesis, Professor Ilpo Koskinen and Dr. Mor Naaman, who gave me excellent suggestions for improvement.

This thesis has been part of the work done by tens of people in different projects and organizations, especially the work and contribution of all the eleven co-authors of the original articles of this thesis. The collaboration with Futurice has been the central part of my doctoral work, and without Hanno Nevanlinna, Mikko Viikari, and Juha Pesonen none of the systems and the related research in this thesis would exist. Not only have they been involved in each of the systems but they have brought their wisdom, ideas, and know-how into the academic issues as well. I would also like to give special thanks to my colleagues Antti Oulasvirta and Mikael Johnson for the constructive feedback and suggestions for improvement of this thesis. I am grateful also to my colleagues, other researchers, and experts who I have had the opportunity to share ideas with, and to people who have made this work possible and fun: thank you Asta Bäck, Minna Forsell, Katri Grenman, Erick Herrarte, Fernando Herrera, Herkko Hietanen, Kai Huotari, Markus Huttunen, Giulio Jacucci, Sami Jokela, Atte Järvinen, Mikko Kiesilä, Vesa Kantola, Harri Kauhanen, Kai Kuikkaniemi, Sari Kujala, Esko Kurvinen, Vili Lehdonvirta, Leevi Lemmetty, Aino Mensonen, Pirkko Miettunen, Samu Myllykangas, Jaana Näsänen, Ville Oksanen, Olli Pitkänen, Mika Raento, Matti Rantanen, Pekka Rehtijärvi, Tommo Reti, Ken Rimey,

Päivi Saarinen, Antti Salovaara, Anu Seisto, Lassi Seppälä, Aura Soininen, Yuri Takhteyev, Sauli Tiitta, Heidi Tikka, Seija Ulkuniemi, Sari Vainikainen, Sami Vihavainen, Mikko Vihonen, Perttu Virtanen, Janne Vuorenmaa, Petri Vuorimaa, Mikko Välimäki, Anita Wilhelm, and Susanna Ånäs.

Finally I thank my friends and family for showing me that there are so much more important things in life than a doctoral thesis. Special thanks to my wife Tuula for both encouraging me in my work and suggesting when not to work, and to my son Antti for being himself. Finally, I would like to dedicate this book to my dear parents Lea and Jukka.

In Espoo, November 21, 2006,

Risto Sarvas

INTRODUCTION

Forty percent of Western European households, in the year 2005, had a digital camera, and estimated seventy million digital cameras were shipped the same year [41]. Other estimates say that almost 300 million camera phones were sold in 2005 [15]. These numbers are an indicator of how media creation devices (still cameras and video cameras) are becoming household items in the industrial countries. Also, the integration of a camera into a mobile phone means that these media creation devices are becoming more personal rather than the single film cameras shared among a family. Taking a snapshot photograph has never before been easier, cheaper, and taken so little physical space. The result is that people are taking more and more photographs and there is no sign of this phenomena slowing down.

Alongside the popularity of digital snapshot photography the households are becoming networked in the Internet through broadband PC access and mobile phones. This ubiquitous connectivity has made it easier to communicate and keep in touch with other people by means of information technology.

This dissertation is written in the middle of these technological changes. In the year 2006 digital cameras have already become a household item in many homes and the sales are still growing. Mobile phones with more advanced hardware, software, and media capabilities, and faster network bandwidth, are marketed as regular consumer phones. The number of digital photographs on people's PCs is probably measured in hundreds or in thousands – and it is growing incrementally because nothing is deleted much due to the decreasing prices of digital storage space.

Therefore, the future vision in this dissertation work is that in the next ten years

almost every Western household will have a digital camera and several mobile phones (smart phones) with photography and video capabilities. These devices will be connected to the Internet. People have tens or hundreds of thousands of digital photographs and video clips taken by themselves or by their acquaintances. Digital media created by people themselves is an everyday part of communication and socializing. And lastly, people have problems managing all the personal media they create and receive.

This dissertation is about what information people would find meaningful and useful in their digital photographs. The underlying belief is that with information associated to media, it is easier to manage, it is possible to create new media, and the storage of the information has value as such. The main media creation device in this work is the mobile phone, which has certain advantages over digital cameras. However, the perspective of this work is digital snapshot photography in general, not a specific device. The objective of this dissertation is to describe and analyze the work done on snapshot photography and metadata to contribute to research and design in media technology.

The word "social" is used throughout this dissertation to describe uses and phenomena that involve interaction among people. Also, many key terms used to emphasize the interaction and group activity have the word social in them. For example, by "social use of photographs" we mean using photographs as means for socializing. By "social activity" in relation to photographs we mean interaction such as sharing photos and commenting photos. By "social awareness" we mean awareness of interaction or social activity, for example, becoming aware of new comments or photos in a photo sharing service.

This introduction chapter first discusses snapshot photography from a historical perspective, how technological advances have been critical in shaping it, and after that, what is metadata and its role in media organization. At the end of this chapter, the research problem, objective, methodology, scope and contribution are stated. The second chapter goes through related work in the area of organizing, sharing, and managing personal photographs. The third chapter describes the work done on

creating media metadata on a mobile phone with the MMM-1 system. The fourth chapter describes the mobile photo sharing system MobShare which was designed to support traditional snapshot photography. The fifth chapter goes through the PhotosToFriends system by discussing the meaning of social activity in photo sharing and the design changes dictated by commercialization. The fifth chapter pulls together the results and analysis of each system for discussion. The sixth chapter concludes the dissertation by discussing the contributions and implications of this work for metadata in snapshot photography.

FROM SNAPSHOT PHOTOGRAPHY TO SOCIAL MEDIA

Media, in this dissertation, means different media types, namely text, audio, image, video, and multimedia. Multimedia is media that is a combination of different types of media, such as a DVD that contains a movie and pictures, a link to a web page, or an interactive media (*e.g.*, a computer game). Whether the media discussed is in digital or analog format should be obvious from the context of the writing.

The media studied in this work is mainly digital photographs. Within the domain of photography our focus is on self-created photos and images, in other words, *snapshot photography*[1] in the digital age. The term snapshot photography comes from the way the photos are captured (easily, instantaneously, and simply, by a single click) and the intentions the photographer has for the photo: no artistic nor commercial intentions, rather photos taken with simple cameras by non-experts for personal use [16, 26, 84]. Snapshot photography is used often to distinguish between professional photography, and especially, amateur photography, where the intentions are more creative and artistic, and the technical skills of the photographer are emphasized.

As current consumer media creation devices include video cameras and other forms of media are being used similar way to the use of photographs, we bundle all

[1] *Näppäilykuvaus* in Finnish.

these forms of media under the term *snapshot media*. Our definition of snapshot media includes video clips, text messages, multimedia shows, and so on, which is created by non-professionals for non-commercial purposes, and the intended audience for the media is themselves and/or their social network (*i.e.*, family, friends, colleagues, etc.).

The term *domestic photography*[2] is often used also for the type of personal photos people have in their homes. Domestic photography includes also photos taken by professionals in studios (*e.g.*, wedding photos or graduation photos) [84].

In other words, the definition of snapshot media is a definition by authorship and the key is that the intended user of the media (*i.e.*, the intended audience) has a personal relationship with the creator of the media (*e.g.*, a friend, a relative, an acquaintance, or as most often is the case: the user is the creator). Therefore, the user has a different kind of relationship to the media than to amateur or professional media.

Amateur media often imitates professional media, and as in the case of *user-created content* or *user-generated content*, they can be modifications of professional media or additions to that. User-created game content is often made by the fans of the game for other fans (or any consumers) of the game to enjoy, and often without any commercial intentions [55]. User-created content is amateur media rather than snapshot media, because creating computer game content (*i.e.*, audio, graphics, or program code) is often non-professional activity but requires technical or artistic skills, and it is not created in a snap but through varying amounts of work and collaboration.

Professional media we define as media created for commercial purposes and often for public audiences. Professional media can be non-commercial such as public and free media archives, however, this media is most often created by people who would be called professionals.

---

[2] *Perhevalokuvaus* in Finnish.

However, the distinctions between snapshot, amateur, and professional media are not unambiguous. For example, a snapshot photograph in today's world may have been captured for a more public appeal. The photographs published in services such as MySpace or IRC-galleria can be defined as snapshot photos: they are mostly taken with consumer cameras and there are no commercial intentions. However, the photos published in these services are public and uploaded into the service just because of that. Therefore, the intended audience (*i.e.*, the users of the media) does not necessarily have a personal relationship with the creator or know who the creator is. These types of photos or media are closer to amateur media, but in contrast to amateur media, they do necessarily try to imitate professional media: they are simply snapshots for public use.

TECHNOLOGY IN THE HISTORY OF SNAPSHOT PHOTOGRAPHY

The history of photography is tied to the technological changes in cameras, film materials, the development process, and the physical images. Technology has played a key part in shaping photography and photos alongside the social norms, conventions, values, visual history, and culture. Snapshot photography itself is a result of technological advancements: technical inventions effectively decreased the costs in photography and made the whole photo creation process easier for the non-professional photographer.

There are four technological milestones in consumer photography. First, the cost of having one's photo taken decreased in the 1850s due to cheaper paper photos and advancements in photo development technology [84]. Owning a photograph of one's self was no longer a luxury product. Second, in 1888 George Eastman's Kodak cameras brought photography within the price range and ease-of-use acceptable for non-professional photographers and effectively started the genre of snapshot photography [16, 26, 84]. A third major step was the technological innovations made in the 1930s that were further developed and extended after the Second World War: automatic exposure, miniaturization, color processes, flash, drop-in cartridge loading, and instant picture cameras [16, 84]. According to Coe and Gates [16], the real cost

of an inexpensive consumer camera in 1939, allowing for inflation, was almost one percent of the first Kodak in 1888, thus bringing photography available to almost all levels of society. The fourth major step is the digitalization of snapshot photography, which is discussed in more detail later on.

One objective of this work is to show how design and implementation decisions made in building technology for snapshot photography affect what people photograph and what they do with the photos. This is by no means a novel insight. Technology has molded photography in the past century and a half and it can be narrowed down to concrete examples. For example, advances in camera lenses changed the distance between the photographer and the object, flash made it possible to take pictures indoors, and automation of light measurement and aperture and shutter settings enabled people to take photos more quickly, easily, and spontaneously. Also, the physical photo itself has changed. The round and oval shape of photos was replaced by the rectangular shape as lenses got better, better film quality has made it possible to make larger photos, color technology has almost replaced black and white photos in snapshot photography, and better lens and camera technology has made possible field of depth, as well as, sharper and clearer photos. [16, 26, 84]

Snapshot photography with its conventions, needs, uses, and culture has also affected photograph technology. Although consumer needs are often emphasized or fully created by marketing, they have had a key role in shaping the technology. The primary end-user force driving photograph technology has been people's need to take and own pictures. This need to get pictures of one's own self and family has its roots in portrait paintings and it has been the seed in turning photography from a small service provided by professionals in the 1850s into a worldwide consumer business it is today [84]. The business created by people's willingness to get their photos taken, and to take their own photos, has further increased the pressure to decrease camera, film, and development costs. As mentioned above, alongside the lowering costs involved, the automating and simplification of the capture and development processes has also been a goal for technology.

Figure 1. A Kodak advertisement from 1905 emphasizing simplicity and ease-of-use, and indirectly encouraging photographers to take the camera outdoors and to travels to foreign countries[3].

A third major drive in camera technology has been the need to make the camera more portable, in other words, lighter and smaller. Already in the early 1900s the marketing encouraged people to take the camera outside the house, which turned out to be a successful proposition making holiday and other outdoor photos one of the most popular sub-genres in snapshot photography (see Figure 1). This portability requirement has created the pocket-size snapshot camera, where the photographer

[3] Picture downloaded from The Ellis Collection of Kodakiana (http://scriptorium.lib.duke.edu/eaa/kodak.html).

has only to point the camera and push a button – shutter speed, aperture, and focusing are all made invisible to the user by automation.

The history of photography can be seen as an interaction between the technological inventions and the end-user needs, whether created by marketing or by people's existing needs. In this research we take the view that this dialog between technology, business, and end-use has molded photography into what it is today. By the end of the 20th century, snapshot film photography had stabilized into the familiar process and business: buying a new film roll and loading it into the camera, taking pictures until the film roll is full, taking the roll to be developed, and then, after either few hours or days, getting the paper pictures and negatives. The key characteristics of this process are: the pictures can not be seen until they are developed, the cost of film and development practically makes people take full film rolls before development, and the pictures are most often paper photos. The fourth major milestone in snapshot photography, the digitalization of photos, is disrupting this familiar process and the whole film-based consumer photography business.

DIGITAL TECHNOLOGY CHANGING SNAPSHOT PHOTOGRAPHY

One of the biggest changes in photograph technology and business has happened in the digitalization of pictures, especially by consumer digital cameras in the 1990s. The digitalization of photographs has changed, and is still changing, people's habits and traditions in capturing, sharing, viewing, re-using, and archiving photographs. In addition to the changes caused by photo digitalization, another major change is the wide availability of Internet access. Photographs are traditionally shown to other people (whether snapshots at home or artistic photos in an art gallery) and the Internet has changed this by making it possible to untie the showing and viewing of photographs from location. Having cameras and network access in mobile phones has emphasized this even more by enabling remote sharing to the Internet immediately after taking a picture, thus skipping the process of transferring the picture from the camera to an intermediary computer before sharing. Below we discuss in more detail these three technological changes in the light of this

dissertation.

Digitalization

The digitalization of photography means that once a photo is taken it is stored in digital format on a separate memory. Lev Manovich calls digitized media *new media*, in contrast to old analog media, and he lists five principles that are the result of the digitization [58]:

1. *Numerical Representation*. Digital media can be described formally and is, therefore, subject to manipulation. Media becomes programmable.

2. *Modularity*. Digital media consists of independent elements which can be assembled to form other media objects without any of the elements losing their independence.

3. *Automation*. The numerical representation of media and the modular structure allow for the automation of operations involved in media creation, manipulation, and access.

4. *Variability*. Another consequence of numerical representation and modularity is that a new media object is not something fixed once and for all. It can exist in different, potentially infinite versions.

5. *Transcoding*. On the level of representation, digital media belongs to human culture. On another level, digital media is a computer object that belongs to computer's own cosmogony rather than to human culture. The computerization of culture by digital media gradually transcodes the cultural categories and concepts into something closer to the world of computer's.

The main effect of having the media as a numerical representation in snapshot photography is that there are no costs directly associated with taking a photograph. The photograph is created in numerical format which only uses re-usable digital memory. The digital format has also made it possible for the photographer to see the captured photo immediately after capture on the camera's screen: there is no wait

period between photo capture and seeing the photo.

The numerical representation, modular structure, and automation have made the process of editing photographs easier and more automatic. Already immediately after capture digital cameras automatically run algorithms on the captured photo to enhance its quality or to encode it to a certain format. Also, the stored photos can be transferred to a computer where they can be further edited, cropped, reshaped, resized, etc. with a wide variety of software ranging from professional tools to automatic programs. The automation can extend also to the way photos are organized and grouped on the camera or the computer.

The variability of digital media can be seen in the way photos are shared, transferred, and re-used. The numerical format of digital media makes it possible to transfer media all over the Internet without any loss of quality. This has made it possible to make photographs potentially visible to anyone connected to the Internet, and to make potentially infinite number of copies of the media. The variability of media is also seen in the ways digital snapshot photographs are re-used and assembled into new media objects: people create photo slideshows, home videos, web pages, photo collages, and picture books from their photos.

The transcoding of culture to follow the world of computers can be seen, for example, in the way we speak: it is not possible to *give* digital photos unless they are on some physical medium, therefore people talk about *sharing* or *sending* digital photos. Also, the storage medium for photographs is no longer measured in the number of photos it can hold but the amount of bytes.

The digitalization of media and the automation and variability associated with it have also made possible effective archival and organization of personal photographs. The management of personal media is, however, not trivial. The number of digital photographs people have is currently measured in hundreds or thousands, and this number is prone to increase over the years, as people accumulate their own photos, receive photos from their acquaintances, and other media forms are created more easily (*e.g.*, video). Organization of personal media, such as photographs, is an investment to the future: the work put into organization pays off when the organized

collection is used to find something. Therefore, the rewards from the organization effort are not immediate, and this is probably one of the reasons why the organization of traditional paper photos is often postponed infinitely. Automation of the organization, on the other hand, faces the problem of personal media being personal in nature. In other words, the information associated with media such as photographs is highly semantic and related to the context of everyday life of each individual: people, relationships, events, stories, family history, feelings, and so on. This personal media management problem is in the core of this dissertation work and is discussed in detail below.

Network Access

The worldwide Internet usage, according to The World Internet Usage Stats [88], is 15.7% of the world population. In North America the usage is 68.6%, in Europe 36.1%, in Oceania/Australia 52.6%, and in Asia 9.9%. Worldwide, the usage has grown 183.4% from the year 2000 to 2005. In other words, a significant number of people use the Internet and the number is growing.

The digital format of photographs has enabled people to take advantage of the computer networks in photography. As mentioned above, this has, first of all, made an everyday practice of showing and viewing of photographs independent of time and location (*i.e.*, the same photo can be looked at the same time in different locations, and at different times at different locations, for example, with e-mail). Secondly, with worldwide Internet coverage, a snapshot photo can be published for anyone to see. The network, therefore, has made it possible to show and view photographs in new ways, and to audiences impossible with film technology. An example of this is the phenomena of *blogging*, where people keep their own diaries or publish other writings of their own. Also, the digital photograph has become a supplement to any form of communication in the Internet. Although no data is currently available, it is a relatively safe assumption that digital images are used in personal and public communication more than paper or other film-based photos were. For example, it is possible to attach images to web pages, e-mails, and phone short messages (MMS).

Mobile Phone Cameras

Mobile phone cameras are currently the most sold cameras in the world [60, 62]. New phone models have an integrated camera and this seems to be a rule rather than an exception. Of the current available Nokia mobile phone models listed on the Nokia product list, 37 out of 62 have an integrated camera[4].

In addition to the camera, the processing power and faster networks enable the consumption of richer media on mobile phones, such as pictures, sound, and video. Therefore, the mobile phone has become a viewing, editing, and storage device for personal media as well a creation device. The key feature of camera phones is the inherent network access which enables transfer of media from the device directly to the network. This feature is used for sharing photos, messaging with media, and publishing media on the Internet. The camera phone is also a ubiquitous device because it is practically always carried along. This "always with" characteristic has inspired people to take pictures in situations where they previously have not had a camera with them [II, 69]. The last feature that makes camera phones a special device is their access to information. Unlike digital cameras, camera phones are computers that have open programming interfaces, which provide access to information gathered and stored on the phone as well as the Internet. This information can range from call logs to near-by Bluetooth devices in addition to information on the Internet (see, *e.g.*, [61, 71]).

A Change in Photography Infrastructure

The digitalization of photographs has made photography part of the ICT (information and communications technology) infrastructure. By infrastructure we mean not only the devices, tools, standards, but also the actions, conventions, and culture around them [79]. From the consumer's or snapshot photographer's point of view this has significant implications. The film based infrastructure was defined by

[4] Nokia website referenced 9.5.2006
http://www.europe.nokia.com/phones/comparephones/compare.jsp?location=FI&language=FI

the following process: take a photo with the film camera, develop the film when it is full, and get paper photos in few hours or days. The photos were often 10x15 paper photos, which could be shown to guests, given as gifts, sent by letter, or even cut to make a scrapbook. As the pictures became digital, this photo infrastructure became a part of the ICT infrastructure previously familiar with personal computers. The ICT infrastructure is more complex than the old photo infrastructure. The number of options for how to look, share, give, show, edit, and store photos has grown hugely. There is new technology to learn (*e.g.*, to send a picture the user needs to know emailing), new mental concepts to understand (*e.g.*, there can be infinite copies and versions of one photo), new billing models to grasp (*e.g.*, in sending photos from one's phone the user is charged by bits not by pictures), and new social norms to fit into (*e.g.*, where can people take photos, what is appropriate). An indicator of this change in infrastructure is the fact that digital cameras are no longer sold solely in camera stores but also in computer stores next to other information technology.

NEW PHENOMENA IN SNAPSHOT PHOTOGRAPHY

Ulkuniemi [84] categorizes the uses of traditional snapshot photos into four functions (her categorization is based on Musello's work [64]):

- *Documentation*, where the purpose is to store events and important people in family history.
- *Connecting*, where social relationships, togetherness, and shared values are strengthened by photographs.
- *Interaction*, such as the interaction at photo capture, the implicit message that the people being photographed are important to the photographer, and the way how photos often start or facilitate conversations.
- *Creation of self-concept*, such as a channel for creativity or to enhance experiences, and to create one's own visual history.

The changes in snapshot photography caused by digitalization of photos and the availability of Internet access are still ongoing, and it is difficult at this moment to guess the final impact of these technologies on the kind of functions as listed above. However, we identify three new phenomena in snapshot photography that are relevant to this dissertation: a public audience for snapshot photographs, a new genre of utilitarian and practical use for snapshot photographs, and third, the use of keyword tagging of published snapshot photos. All of these phenomena can be seen as snapshot photography either based on the definition of functions above or from the characteristics of snapshots (*i.e.*, the media is captured by simple consumer devices and not for professional or artistic use).

Public Audience

Due to the physical format of snapshot photographs, the audiences for them have been limited. Mostly they have been people the owner of the photos is familiar with: family, relatives, colleagues, friends, and so on. As mentioned above, with digitalization and the Internet there is no longer this limitation. Home pages on the World Wide Web can be seen as one form of reaching a larger audience than mere acquaintances. However, web pages can be lost in the Internet and it can be the case that the only people visiting the public home page are the same acquaintances as before.

In the last years web services such as MySpace[5] in the U.S. or IRC-galleria[6] in Finland have gained significant popularity by providing a central place and a service for people to have their electronic profiles published and linked with each other. These sites are often called "social networking websites". Rather than hiding behind a virtual identity, the users present themselves as they are. For example, in the IRC-galleria service the users have to give their social security number and a photo of themselves in the registration process. According to Alexa Internet [2], MySpace is the fifth most popular web site in the world, and IRC-galleria is the 4th most popular

[5] http://www.myspace.com/

[6] http://www.irc-galleria.net/

web site in Finland. According to the IRC-galleria web site, about half of the Finnish teenage population uses the site.

Photographs taken by the users and posted on their web pages in the service play a key role. In addition to listing personal information, hobbies, thoughts, favorite music, and so on, the photos are used to build a certain image of the person inside the community. This image depends on the motives of the user to use the service. The motives seem to range from simply having a presence to becoming the most popular person within the service. Also, the profiles are used to maintain existing social relationships, as well as to form new ones with people outside the physical social network. For example, the IRC-galleria service is often used as a "business card" in instant messaging conversations.

The photos are taken with regular consumer cameras (digital cameras or camera phones) by the people themselves. The photos are not taken for commercial purposes and not for aesthetic appeal as such. When compared to the functions of snapshot photography listed above, the uses and motives for photographs in these social networking services do not differ from the traditional ones: the pictures published in the profiles are for documenting the person's life, to connect people and form relationships, to facilitate interaction, and to create a self-concept or an image of the person. However, the audience for these photos is present in all of the functions: the documenting, connecting, interaction, and creating of self-concept are all made with the public nature of the service in mind. The new phenomena these services show is that snapshot photos are used as effective means for communicating and socializing with known people or with previously unknown people.

Utilitarian Photos

The ubiquitous nature of camera phones and the zero cost of taking photographs have created a new genre in snapshot photography: utilitarian photographs. These photos are taken as reminders, to store information, or to demonstrate something visually. Examples of these in our research were copying product information by taking a photo, taking a picture of a house to show to an architect, and taking a picture of a car for sale to a friend who was looking for used

cars[7]. The characteristics of snapshot photography are present in these photos (*i.e.*, consumer cameras, and non-commercial and non-artistic motivations), but utilitarian photos do not fit the four categories described above (although one could argue that the four functions are always implicitly present in any snapshot photo). Therefore, it seems that in addition to the traditional categories or genres of snapshot photography (*i.e.*, people, leisure, events, travels, hobbies [16, 84]) the new mobile phone camera which is "always with" has created the utilitarian photograph.

Tagging Free Keywords

The third phenomenon listed here is related to metadata. Metadata, especially metadata for photographs, is discussed in more detail in the following chapters, but here it is suffice to say that it is information associated with a picture, which is stored with the picture or linked to the picture. Tagging, in this context, means adding textual descriptions to a picture, and these descriptions are often one word (*i.e.*, a keyword) and the user has practically full freedom to use whatever word or words (*i.e.*, free keywords without any pre-defined vocabularies or structures). This is not a novel idea or activity, adding free keywords to media objects to facilitate future searches is a widely accepted convention. However, the new phenomenon in tagging with free keywords is that it is actively done as part of social networking and also outside any professional or organizational context. Examples of tagging services are MyWeb 2.0[8] and del.icio.us[9].

The prime example of tagging of photographs is the Flickr[10] service. In November 2005, Flickr reported 1.5 million users and over 60 million photographs, out of which roughly eighty percent were public [75]. The key functionality of Flickr is the tagging: anyone can create a tag and associate it with any public photo.

---

[7] More examples can be found in a competition by a Finnish magazine *MikroPC* on innovative uses for mobile phone photos [46].

[8] http://myweb2.search.yahoo.com/

[9] http://del.icio.us/

[10] http://www.flickr.com/

Therefore, tags are not pre-organized in any way. For example, the most popular tags (31.1.2006) seem mostly to be about the content of the photos: the location ("japan", "home"), the object ("me", "cat"), and the event ("party", "vacation"). But not all are about the content: "cameraphone" is mostly in pictures taken with a camera phone (although some are pictures of camera phones), or "bw" in black and white pictures. Some of them are ambiguous, for example, "family" can mean pictures of family or pictures for family to view.

The tagging in Flickr seems to have three general purposes: the organization of one's own photos by tags, browsing of all the public photos in the service by clicking on tags, and to form communities of interest around a group, which is, in practice, a special kind of tag. The photos in Flickr range from snapshot photos to amateur or professional photos. The new phenomenon of having a public audience for snapshot photos is also strongly present in Flickr. However, the tagging functionality seems to play a key role in making one's photo appeal to the public: tags are the means to attract viewers or to find people and form new relationships.

The motives for adding information to personal photographs are relatively low, as it is discussed later on. The main reason for people not to annotate their photographs is that there has been little benefit and they have seen no reason to write down information which they will always remember. The tagging in Flickr is an exception to this. Our hypothesis is that the role of tags in socializing, communicating, and making one's photos more attractive are the reason why the tags in Flickr are widely used. The new phenomena in snapshot photography is adding information into photographs as means for communication or socializing, or to attract new people to form an interest in the photo, especially beyond one's traditional social networks. It would seem that in addition to the public photos in Flickr, the tags themselves have the four functions listed by Ulkuniemi above.

SUMMARY

The examples of public uses of snapshot photographs described above get

public attention, press coverage, and are accessible to study because they are public. For practical reasons it is more difficult to observe the use of Flickr or IRC-galleria for private purposes (*i.e.*, photographs or comments that are shared within a closed group). However, private use is the traditional use of snapshot photographs: the photos are shared only within limited and known groups of people, and the owner of the photos has control over who can see the pictures. This tradition is much due to the physical form of traditional photos (*i.e.*, paper photos and slides). Nevertheless, current habits, traditions, uses, attitudes, culture etc. come from having photographs in physical form, and these "soft" characteristics of snapshot photography do not change as quickly as the "hard" technology. The scope of this dissertation is the traditional, more private use of snapshot photographs. However, we acknowledge that people's perceptions on privacy of photographs are changing, and people are more than willing to exchange some of their private snapshot photographs to gain popularity or to form new relationships, as the examples above discuss. Also, the transitions from a snapshot photographer to an amateur or a professional one are easier in services that do not make those distinctions (*e.g.*, Flickr). Thirdly, we are aware of new genres and uses in snapshot photography, such as utilitarian photos.

The photo services described above have also a technical feature in common: due to the central architecture (*i.e.*, a web server where the media is stored and which is accessed via a web browser) all the pictures, comments, tags, user information and history, links between users etc. are in one place. In other words, there is lot of user-created information linked to the photographs, which suggests that perhaps some of the information could be stored with the photograph as metadata. This dissertation looks into the social activity around photographs captured in similar services and what of that activity would make sense to store as metadata. The following section goes through the existing media metadata solutions for snapshot photography and the rationale behind the metadata approach in self-created media.

The objective of this chapter was to demonstrate that snapshot photography is at its fourth milestone due to digital technology, and this dissertation work contributes to the new technological changes and the new phenomena happening in snapshot photography. Broadly speaking, the technology focused on is media

metadata in central services, and the phenomena is socializing and communicating with photographs.

METADATA

As discussed above, the media management problem facing consumer (*i.e.*, snapshot) photography is the organization and management of the increasing number of pictures taken. This problem has manifested itself partly because people tend to have more than one camera, and mainly because digital technology lets people take as many pictures as they want without any development costs of film. How could personal photos be organized so that they could be easily searched and browsed for the vast variety of old and new uses people have for photographs (or more broadly: media)? The most prominent answer to the problem is the use of metadata. This is because the same problem of media management has already been addressed with metadata solutions in public archives and in commercial media companies. Through these solutions there is a vast variety of standardization done to describe digital media contents, rights, and other information for organization and search purposes.

This chapter discusses what metadata is and how it is understood in this research. The focus of this research is media metadata, and also, how mobile phones as media creation devices provide new metadata generation technology. At the end of this chapter we also discuss the metadata-related practices people already have with their paper photos.

Metadata has its origins in libraries and in library science. Libraries have the longest tradition of organizing information for the purpose of retrieval and posterity [81, p6]. Simply put, to organize library material for searching it was more convenient to make a catalogue of library cards than search the library by going through the actual books and other material. The information stored in library cards reflected the search needs that people had and still have: author, title, year, publisher, physical location in the library, and so on. This information is standardized and the standard

descriptions for library objects follow such standards as MARC (Machine Readable Cataloging Record), ISBD (International Standard Bibliographic Description) or AACR2r (Anglo-American Cataloguing Rules, Second Edition, 1988 revision). The major research goal of Library and Information Science (LIS) is to organize knowledge and recorded information for keeping a usable record of human endeavors for posterity [81, p.2]. Therefore, the approaches and lessons learned from LIS promise solutions to the snapshot photographer's media management problem. A general and widely used example of the LIS approach in organizing digital objects (*e.g.*, digital media) is the Dublin Core Metadata Initiative. The Dublin Core Metadata Element Set is listed in Table 1.

Table 1. The Dublin Core Metadata Element Set [38].

1. Title	9. Format
2. Creator	10. Identifier
3. Subject	11. Source
4. Description	12. Language
5. Publisher	13. Relation
6. Contributor	14. Coverage
7. Date	15. Rights
8. Type	

Metadata is literally "about something that is a given" (*meta* is a Greek prefix meaning "about" and *data* is the plural of the Latin word *datum* "something given"), or more simply: "about data". The somewhat abstract literal meaning of the word already suggests that it can mean a myriad of things. Therefore, there is no one unambiguous definition of what is metadata. In the information technology practice and sciences it can be narrowed slightly to mean digital information that describes a set of digital data, which can be a digital object. The data which is described can be metadata itself, and therefore, there is no sense in defining a level of abstraction for metadata. Examples of metadata in information technology include metadata about files (*e.g.*, the size of the file, the date of its creation), metadata about programs (*e.g.*,

version numbers), and metadata about media objects (*e.g.*, author, owner, or contents).

The term *ontology* has also several different meanings – even within information technology and sciences. An ontology, in this dissertation, is the structure of the metadata which is lexical (*i.e.*, a vocabulary or a dictionary). The ontology is the hierarchies, categories, and interconnections between metadata. For example, a photograph may have hierarchical metadata "people → family → brother" where the ontology defines that "brother" is "family" who are "people".

What is often not emphasized or is altogether neglected in the definitions of metadata and ontologies is that both the metadata and its structure are a *limited* amount of information, and therefore, a *selected* set of information. For metadata to be computable and understandable, its format and structure has to be defined. It can simply be a single, non-hierarchical field for any kind of information, but nevertheless it has to be defined to a certain degree for other people and computers to find and use it. This means that there is a *purpose* for the metadata that the ontology reflects. To put it simply, one could ask the following question about all metadata and ontologies: from all the possible information that could be described in the metadata and the ontology, why did someone choose these pieces of information in this kind of a structure.

For example, mp3 audio files most often use ID3 metadata standard to describe the audio file's song title, artist, album, year, and genre [40]. The ID3 standard also includes a place for comments. The ID3 (version 1, *i.e.*, ID3v1) metadata reflects its designer's vision of what the metadata will be used for: describing one song that has one artist and a year, and the song belongs to an album and a genre. This metadata would not suit well, for example, an mp3 file of one family's children speaking recorded over the years.

In addition to the selections done by the creators of metadata standards and ontologies, Bowker and Star [10, p107-108] describe the following as the barriers in any knowledge system or encyclopedia:

- *Data entry as work*. Entering data always requires resources and personnel; it will inevitably contain mistakes and cultural variations in interpretation and omissions.

- *Convergence between the medium and the message*. The technology used for storing information is always limited, and what gets stored is at best what can be stored with available technology. The encyclopedia comes to mirror the affordances of its technological base.

- *Infrastructural routines as conceptual problems*. No knowledge system exists in a vacuum, and the work of ensuring compatibility is not only technical work, it challenges the very integrity of any unifying scheme.

Another implicit characteristic of metadata and the ontology, is that it should be objective and universal. Again, if we look to library card catalogues as an example, the information stored in library cards aims to be as unambiguous and objective as possible. This objectiveness is due to the public nature of library metadata, meaning that the information should be understandable to almost anyone who knows the language (*i.e.*., metadata described in Finnish is not understandable to non-Finnish speakers although it could be objective). This objectiveness is emphasized in the standards for metadata which by definition of a standard try to be objective and understandable (see, *e.g.*, Table 1). Of course, the contents of metadata can be subjective. For example, the comment metadata in the ID3v1 standard could include a subjective comment by someone and would still be within the spirit of the standard (it is actually not defined what the comment field in ID3v1 should contain). Nevertheless, the metadata field (*i.e.*, the category label) "comment" implies that this piece of information can be subjective, and therefore, it is understandable to a public audience that the information there can be whatever.

A DEFINITION OF METADATA

Metadata in this dissertation means information about a digital object or a set of digital objects. Objects in this work are digital media objects, mainly pictures.

Therefore, metadata can be information about a photograph or information about a collection of photographs.

Metadata is often defined to be either *descriptive metadata* or *semantic metadata*. Descriptive metadata is information that is external to the meaning of the document [3, p.143], such as the information in the Dublin Core Metadata Element mentioned above (see Table 1). Descriptive metadata can be called also *syntactic metadata*: information about the syntax or structure of the object. Semantic metadata is information about the subject matter itself – the semantics of the object. It can be keywords or more sophisticated and structured information about, for example, who are the people in a photograph or what domains of research an academic paper belongs to.

The metadata can be anything that can be described by text, and hence readable, but not necessarily understandable, by humans. Whether the metadata is text or binary code is a matter of interpretation, because all text in digital form is binary code. The metadata should be machine readable, meaning that a computer can be programmed to represent the metadata.

The focus of this work is in *media content metadata* and how it could be applied to snapshot photography. The other kinds of metadata associated with media are briefly touched upon. What is meant by *media* is discussed previously. *Content metadata* is used in this work on information about the contents of the media object (or a set of media objects). By contents we mean information about the subject matter – the semantic information. For example, who are the people in the photograph or where was it taken. The content of, for example, a photograph is closely related to the context of the photography situation. To use the definition of context by Dey and Abowd [22], the primary context relevant to a user and an application is location, identity, time, and activity. Therefore, the context of a photography situation (*i.e.*, interaction between the photographer and the camera) can be the content of the photograph. However, the difference between context and content is that the context of a given situation (*e.g.*, the capture time *T*) is frozen but the semantic contents of a photograph can change over time (*e.g.*, a girlfriend becomes a wife, a tourist location

becomes home town, *etc.*). Therefore, the location where a photograph was taken can be the context of the picture-taking situation and the content of the photograph. Because the term *context* has connotations to context-aware systems and applications that store or leverage the context at a given time, we use the term *content metadata* to emphasize that the information associated can be independent of a system and not tied to any situation.

As the ID3v1 example demonstrates, metadata (and ontologies) is designed for a purpose, and therefore, includes human selection of described information, and as Bowker and Star [10] argue, the data entry work, medium, message, and routines have an effect on the metadata representation scheme.

MEDIA CONTENT METADATA

The idea behind media content metadata is twofold. First, to describe the content of a media object to enable locating of media objects and searching their contents without going through the media. This is because media objects are most often significantly larger in size than metadata descriptions, and therefore, searching only the metadata demands less computational resources. Second, to attach information to the media object that can not be inferred from the media itself but supports common uses for the media, for example, information about the copyrights or creation process.

An example of media content metadata is the MPEG-7 standard. It standardizes the description of multimedia content supporting a wide range of applications. It does not only address the description of the media content but also the organization of the content, navigation and access, content management, and user interaction. The content descriptions in MPEG-7 are further divided into two: the structural aspects of the media (*e.g.*, spatial, temporal, and spatiotemporal properties), and the conceptual aspects of the media (*i.e.*, describing the content from the viewpoint of real-world semantics and conceptual notions). The conceptual aspects of the media, that is the content semantics, can be described in MPEG-7 by specific semantic

entities, such as, objects, relations, agents, time, place, and events. Figure 2 is an example of a semantic description made with semantic entities of MPEG-7. [57]

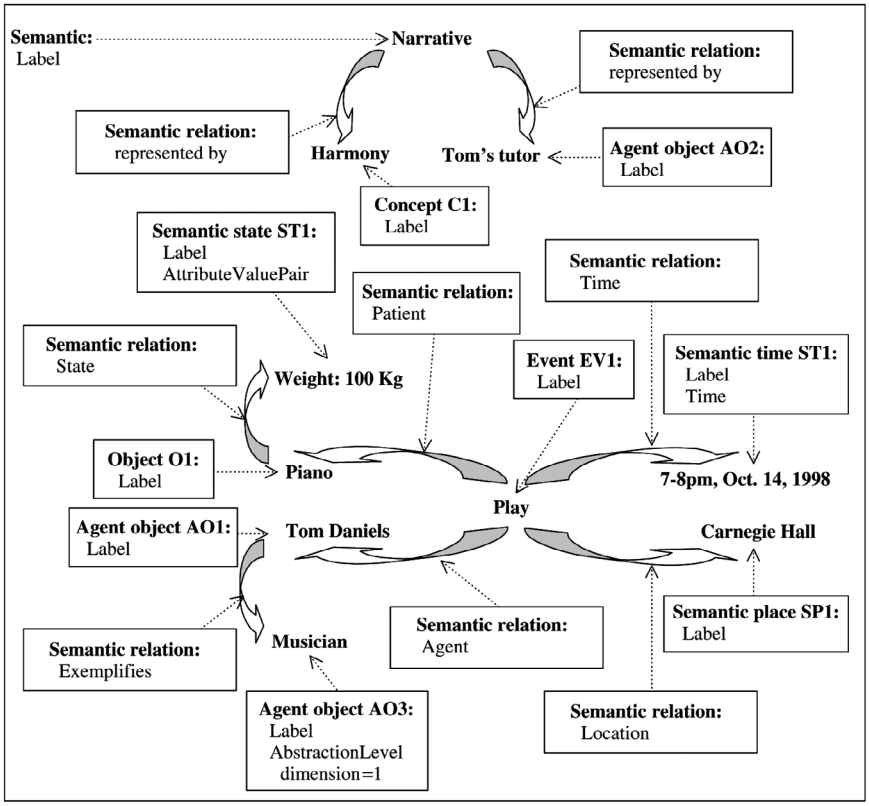

Figure 2. Example of an MPEG-7 semantic description of a piano recital video [57, p132].

For the domain of snapshot photography, MPEG-7 provides a very extensive tool for describing media. However, the MPEG-7 standard, like the MPEG standardization body itself, is focused on commercially produced media or media produced for public use. The selections and design decisions made on the standard reflect this: for a media object there is a producer, a rights owner, and a consumer. For example, the classification preferences for the users of the media are country of origin, date, languages, form, genre, subject, review, and parental guidance [57, p170]. These categories are not intuitive in the domain of snapshot photography, and the kind of semantic descriptions exemplified in Figure 2 create a huge annotation burden for the user, and are very complex.

In software applications designed for organizing snapshot photography some use metadata annotation as a tool for organization. These metadata ontologies are coupled with the application and do not follow any standard. For example, Apple iPhoto uses pre-defined tags or keywords that the user can insert to a single photo or to a set of photos at import (*e.g.*, Favorite, Family, Kids, Vacation, and Birthday). The keywords are then used in searching the iPhoto photo library. The photo organization application Picasa by Google provides the user with a functionality to tag photos with labels, but there are no ready labels or any suggestions on what the labels might be. These labels are then used as one means of searching the picture library. The main organizational information is the folder location on the hard disk and the time and date information of individual photos within a folder.

Adobe Photoshop Album and Adobe Photoshop Elements use metadata annotation as the primary means of organization. The user can tag individual photos with metadata that can be hierarchical. For example, a single tag "Risto" can be in the "Family" category which is under the "People" category (see Figure 3). Before the user starts to tag her photos, Adobe Photoshop Album already has ready-made categories for the tags, namely, People, Places, Event, and Other.

Photo organization applications made in academia study and take more advantage of the possibilities of using metadata to organize personal media. The state of the art and systems related to this work are discussed in detail later. However, the key metadata used in them are parallel to the ones in Apple iPhoto and the Adobe applications: people (who is in the picture), time (when the picture was taken), places (where it was taken), and events (what was the event where the picture was taken). Except for time (see below on EXIF metadata), none of this information is automatically available in pictures taken with current digital cameras.

Figure 3. An example photo and the Adobe Photoshop Album metadata annotated by the user. "Perhe" means family and "Työkaverit" means colleagues.

There are also ontologies for describing photographs in the domain of professional photography and public archives of photographs (*e.g.*, museums, or art galleries). An example of these is the VALO ontology currently under construction in the Helsinki University of Technology TKK [85]. The ontology is based on the General Thesaurus in Finnish (YSA) and on the Finnish thesaurus of photography. VALO is designed for professional photographers, researchers on photography, and for amateur photographers. Examples of the terms in the thesaurus for photography are "abstract photography", "journalism", "print products". Ontologies for professional photography and photo archives are not designed for the snapshot photographers personal use, and therefore, are not usable in that domain.

OTHER MEDIA METADATA

In addition to describing the content of the media, there is media metadata to describe other things considered relevant in using the media. Examples of these are licensing information, production and publishing information, technical information, summarization and navigation information, and information for user interaction [57]. Another form of metadata is user-created information. User-created media metadata can be for other users such as ratings, comments, or links to other media, or the metadata can be about the user herself, such as preferences or user history.

Make - Canon Model - Canon PowerShot S410 Orientation - Top left XResolution - 180 YResolution - 180 ResolutionUnit - Inch DateTime - 2006:03:03 11:53:58 YCbCrPositioning - Centered ExifOffset - 196 ExposureTime - 1/60 seconds FNumber - 2.80 ExifVersion - 0220 DateTimeOriginal - 2006:03:03 11:53:58 DateTimeDigitized - 2006:03:03 11:53:58 ComponentsConfiguration - YCbCr CompressedBitsPerPixel - 3 (bits/pixel) ShutterSpeedValue - 1/60 seconds ApertureValue - F 2.80 ExposureBiasValue - 0.00 MaxApertureValue - F 2.80 MeteringMode - Center weighted average Flash - Not fired, compulsory flash mode FocalLength - 7.41 mm UserComment - FlashPixVersion - 0100 ColorSpace - sRGB ExifImageWidth - 2272 ExifImageHeight - 1704 InteroperabilityOffset - 1860 FocalPlaneXResolution - 8114.29 FocalPlaneYResolution - 8114.29 FocalPlaneResolutionUnit - Inch SensingMethod - One-chip color area sensor FileSource - DSC - Digital still camera CustomRendered - Normal process ExposureMode - Auto WhiteBalance - Auto DigitalZoomRatio - 1.00 x	Maker Note (Vendor): - Macro mode - Macro Self timer - Off Quality - Fine Flash mode - Not fired Sequence mode - Single or Timer Focus mode - Single Image size - Large Easy shooting mode - Manual Digital zoom - None Contrast - Normal Saturation - Normal Sharpness - Normal ISO Value - 200 Metering mode - Center weighted Focus type - Close-Up AF point selected - Exposure mode - Easy shooting Focal length - 237 - 711 mm (32 mm) Flash activity - Flash details - Focus mode 2 - Single White Balance - Auto Sequence number - 0 Flash bias - 0.00 EV Subject Distance - 13 mm Image Type - IMG:PowerShot S410 JPEG Firmware Version - Firmware Version 1.00 Image Number - 1868622 Owner Name - Y

Figure 4. The EXIF metadata stored by Canon PowerShot S410.

Probably the most used and available metadata standard in snapshot photography is the Exchangeable Image File Format (EXIF). Most of current digital cameras and camera phones store EXIF metadata in the image file at the time of capture. EXIF describes the technical information about the photo capture, such as, camera make and model, date and time of capture, width and height of the image in pixels, whether the flash was used, focal length used, CCD width, exposure time, aperture value, and metering mode (see Figure 4). However, as discussed in this dissertation, and especially in [V], it is questionable whether the technical information about the state of the camera is useful in organizing personal media (excluding the time and date information).

PROBLEMS WITH MEDIA METADATA AND ONTOLOGIES

In the last few years there has been criticism towards the metadata approach in organizing media. The ideal situation, from the metadata point of view, is to have the metadata generated at the same time as the object and there would be no new objects without metadata. However, as Bowker and Star argue [10], data entry requires work. Although it can be automated in some cases, the creation of that automation requires work and the automation is limited by the currently available technology. Also, metadata requires maintenance and management, because some information can change and so can metadata standards, ontologies, and the technology. Bulterman, in his article [11], criticizes the metadata approach in the case of media with the following arguments: much of the data gathered through standards like EXIF is irrelevant, metadata entry is prone to human errors, some of the metadata created by applications is useless, and some of the metadata simply does not apply. Bulterman summarizes the issues in metadata creation: it is context-sensitive, culturally biased, and time-variant.

A critical issue in automatic metadata creation is the so-called *semantic gap*. Smeulders et al. [77] define it as follows in the context of image retrieval: "The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation." In other words, it is the gap between the rich meanings that people want when they search and browse media and the shallowness of the content descriptions that can be computed automatically [23].

The alternative approach in organizing and searching information to the metadata approach is the Information Retrieval (IR) approach. Often the metadata approach is called *Data* Retrieval to emphasize the difference to *Information* Retrieval [3]. Data retrieval differs from IR by searching only the metadata associated with the information, therefore, it assumes that the metadata has a well defined structure and semantics. The IR approach does not rely on pre-defined structures and semantics, therefore, it has to interpret the searched content to be able to provide relevant results [3]. For example, searching through text is so efficient with current

technology that it is not necessarily economical to abstract the text into metadata prior to searching and retrieval [25]. A concrete example of a successful commercial IR approach is the Google search engine, which does not rely on the searched content having any metadata at all. On the other hand, as Foote points out, even the Google approach (*i.e.*, not relying on metadata and ontologies) requires some indexing and abstractions to work, therefore, it is also limited to searching only the information in the abstraction [25]. The IR approach does not solve the semantic gap problem. It only moves the problem away from metadata creation to information interpretation.

Another alternative to the metadata and ontology approach that is gaining popularity is the tagging phenomena discussed earlier. Adding free keywords into pictures (Flickr) or to web pages (MyWeb 2.0) can be seen as media metadata. However, these user-created free keywords are not structured, there are no rules on how to name or use tags, and they are not always created with a public audience in mind. Therefore, they are information about the media (*i.e.*, metadata according to its broad definition), but not metadata as a standardized set of classified information and rules of conduct. Ironically, these differences to the traditional characteristics of metadata seem to be the reason behind the popularity of tagging. As the overhead of following rules and ontologies is removed, the data entry work is lessened significantly. The organizational downside of tagging is that the tags are open to interpretation like any free keywords that represent an individual's associations. As mentioned before, the motives behind tagging seem to be the way tags enable communication and socializing with other people, and the organization by tags is a side-product of this.

To summarize, the traditional approach of describing media by metadata is criticized because the ideal of having all relevant and objective information pre-defined does not work in practice. Bowker and Star point out that any knowledge system has to address the following three issues: data entry requires work which is prone to cultural variations, available technology limits what information can be stored, and the requirements for backward and sideways compatibility challenges the whole integrity of the system [10, pp.107-108]. Bulterman argues that creating

metadata is context-sensitive, culturally biased, and time-variant [11]. As the work done in this dissertation demonstrates, these issues are even more problematic in the domain of snapshot photography where the benefits of creating metadata are vague, and the requirements for the information stored highly semantic. Foote suggests that the query, that media retrieval and search should be able to address, is as semantic, personal, and contextual as "What media will best serve me at this moment?" [25].

MOBILE PHONES AND MEDIA CONTENT METADATA

Mobile phones with cameras are unique media creation devices from the perspective of metadata. As mentioned previously, mobile phones by their nature as communication devices have two characteristics that traditional cameras do not have: the inherent network connection and access to contextual and social information. From the point of view of metadata the access to contextual and social information, as well as network resources, open up new opportunities to generate media metadata. For example, the Merkitys-Meaning mobile phone application adds at the point of capture location information (GPS coordinates if available, GSM cell information, country and city), the Bluetooth environment, events from the phone's calendar, and allows the user add tags and descriptions [61]. Using contextual information provided by the phone can be used as an alternative to doing content-based media analysis to infer media content metadata [20]. A more detailed discussion of these opportunities and the work on leveraging contextual information and network resources is discussed later, and also, in [I].

METADATA CREATION AS SNAPSHOT PHOTOGRAPHER ACTIVITY

There has there been information associated with snapshot photographs in the past, even before digitalization and media metadata. The annotation of photographs has been manual writing behind a paper photo, next to a photograph on an album, on the sleeve where the developed photos are received, or in the proverbial or actual shoebox where paper photos are stored. This information has been written for

organization purposes (*e.g.*, "Norway 2005" on the photo sleeve for pictures taken on the family trip to Norway), for storing information considered valuable (*e.g.*, names of all the people in the picture so the information is not forgotten), or for describing the photo (*e.g.*, a photo is sent in a letter to someone, or put in a photo album).

The amount of manual annotation and organization of paper photos varies between individuals. In our user interviews conducted for this dissertation research people showed their personal photo organization and albums. Some people had a habit of putting all their paper photos into albums and writing information next to the photo. Other people only looked at the paper photos as they got them from the developing service and then archived them in some box. Some of them wrote a descriptive word on the photo sleeve. A shared goal for all the interviewed people was to some day find the time and effort to go through the photos and organize them.

The physical nature of paper photos limits the situations where the pictures are shared or showed to other people. A common rationale given by the interviewees for not annotating photos was that they will be always there to tell about the photos, and they will never forget what the event was or who the people were. One interviewee said that if she forgets a person in a picture, then that person was not important. These interviewees were aged 25-26, so their age probably had an influence on their perspectives on personal and family history.

There exists a motivation and some analogy in snapshot photography in relation to metadata. It is not an alien thought for people to associate information with their photos. From this practice of writing information on paper photos or albums rise two important points: first, people already perceive the annotation work as labor-intensive, second, the most thorough annotation is often done as a part of some other task, such as, sending photos to a relative, or making a photo album as a gift. The habit of annotating digital photos can be seen in the way some people rename the image files to have more descriptive names. One of the contributions of this dissertation is to design metadata annotation in such a way that it is a side-product of some task which is motivated, such as sharing photos.

RESEARCH QUESTION

The main research question of this dissertation work is what kind of metadata would support the uses people have for snapshot photographs. This question is divided into further questions, mainly: what are the uses people have for snapshot photographs, and what information in the form of metadata would support those uses? Also, as the method of research is designing and constructing consecutive systems there are research questions on how the design decisions affect the end use, and can the metadata be designed independent of system?

OBJECTIVE OF RESEARCH

The objective of this research is to provide understanding and new knowledge on metadata in snapshot photography from the perspective of the photographer (*i.e.*, the end-user). Part of this objective is to show that snapshot photography is a domain of its own, and it has fundamental differences to public or commercial media domains, and therefore, the metadata designed for these domains does not work in the snapshot domain. To reach this objective, three systems were built and evaluated in user trials. In addition to providing data for designing metadata, the objective of constructing and evaluating these systems was to generate knowledge about constructing online photo sharing systems and knowledge on how and for what purposes people use these systems.

RESEARCH METHODOLOGY

The methodological framework used in this research is from design-science in information systems (IS) research described in [37]. The basic idea of design-science research is that designing information systems and studying users' behavior with it are complementary. Therefore, the fundamental research process is an iterative cycle of building new technology to address identified requirements and of evaluating the impact of the built technology in interactions among people, technology, and

organizations [37].

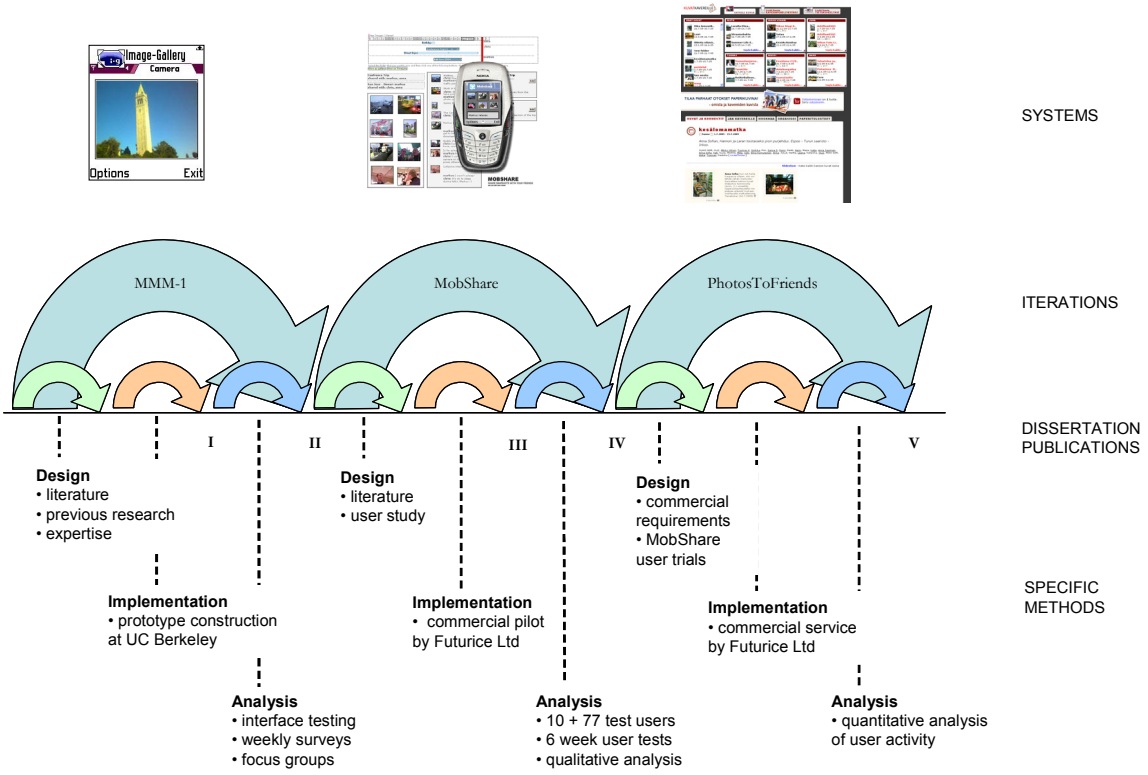

Figure 5. The three systems built in this research, each system consisting of a design, implementation and analysis parts. Also, the original articles of this dissertation are marked to show their focus in the iterations. The methods used are listed below the iterations.

The philosophy adopted in designing the systems and analyzing the results is user-centeredness. In the spirit of Norman and Draper [68], the aim of the technology built in this work and the analysis of it is to serve the user rather than use a specific technology. Therefore, in addition to designing a metadata ontology for users to use [I, II], the alternative approach has been to let user requirements and actual use drive the design and see what metadata could facilitate that use [III, IV, V].

The new technologies built in this research are the three systems described below: MMM-1, MobShare, and PhotosToFriends. The user behavior related to these systems is studied and evaluated with qualitative and quantitative methods to gain a deeper understanding of the issues related to media use and metadata in photo

sharing systems. Therefore, this research can be seen as three consecutive cycles of building technology and evaluating it (see Figure 5).

DESIGN-SCIENCE RESEARCH

The design-science paradigm, according to [37], is fundamentally a problem-solving paradigm which has its roots in engineering. Behavioral science, on the other hand, has its roots in natural science research methods, and in the context of information systems its goal is to develop and justify theories that explain or predict organizational and human phenomena surrounding the analysis, design, implementation, management, and use of information systems. These theories impact the design of further information systems, and that is how the cycle progresses. [37]

In a research field, such as personal media and metadata, where existing theories on user behavior are insufficient the design-science research approach provides a methodological framework. Taking it as a fact that technology and behavior are inseparable in information systems [37] the method used in this dissertation work is to start from current state of the art and construct and study working systems to gain a better understanding. By designing and implementing technology it is possible to study the interactions between behavior and design in photo sharing systems. And by this method it is possible to study and discuss the implications for media metadata in self-created media based on existing technology and empirical user studies.

However, the design-science research described by Hevner in [37] does not fully apply to the research done here. Information systems (IS) research studies systems that are implemented within an organization for the purpose of improving the effectiveness and efficiency of that organization [37]. The organization, in the discipline of IS research, is a business organization, and IS research addresses the interplay among business strategy, IT strategy, organizational infrastructure, and IS infrastructure [37]. The research in this dissertation is not done in a business context, although commercial requirements are considered. The organizations studied in this

work are social networks of acquaintances (*e.g.*, friends, family, and colleagues), the motives and rationale behind user behavior are from the perspective of an individual snapshot photographer, not an employee, and the systems constructed and evaluated are built for either research purposes or as commercial systems. Therefore, the IS design-science research is used more as a methodological *framework* rather than a strict method to follow. Although the objectives of IS research are not the same as in this work, the main approach of studying relatively unknown phenomena by the iterative build and evaluate cycle can be applied.

USER-CENTERED DESIGN

The objective of this work was to ground the system design decisions and base the analysis on principles of user-centered design or user-centeredness. However, there is no agreed upon definition of user-centered design (UCD) or user-centered systems design (UCSD), therefore, to describe the user-centeredness of this dissertation work, we use the twelve key principles for user-centered systems design described by Gulliksen et al. [36]. These principles are based on existing theory and literature [6, 7, 8, 12, 17, 33, 35, 42, 48, 54, 59, 67] and the research and experience of the principles' authors. The objective of comparing these principles to the systems design in this dissertation is to communicate and ground the level and nature of user-centeredness both in the systems as well as the contributions of this work for metadata design.

However, it is good to bear in mind that the principles described by Gulliksen et al. are for industrial software development projects, not for researching new technology and new user behavior in an academic context in projects with less than ten people. For example, the systems in this dissertation were used to study the uses, situations, goals, and tasks of the users, and therefore, for the designers to have a full understanding of these beforehand was not possible. Also, the goals of the three systems varied in their user-centeredness: the first system (MMM-1) was built primarily to test new technology, the second system (MobShare) was designed to understand user behavior, and the third system (PhotosToFriends) was designed to

make a business out of photo sharing activity. In the following, the system design work of these systems is seen as one with three consecutive system iterations because the building of each relied heavily on the experiences and user studies made on the previous system (except, of course, the first system MMM-1). In the following the twelve principles of Gulliksen et al. are in bold typeface.

User focus – the goals of the activity, the work domain or context of use, the users' goals, tasks and needs should early guide the development. Literature on people's use of snapshot photos was studied, and a specific user interview study was made on mobile camera use [69]. Extensive user studies were made on the first two systems, the goal of which was to understand better context of use, users' goals and needs, as well as the impact of the built technology on these.

Active user involvement – representative users should actively participate, early and continuously throughout the entire development process and throughout the system lifecycle. Users were not directly involved in the system design – they did not participate in the project meetings or were part of the design process as members. However, as mentioned above, extensive user studies were made between the individual systems which had a strong influence on design. The user studies were conducted in real user context, not in a laboratory. The systems were also used by acquaintances of the system designers during the development process, and they gave feedback and suggestions to the versions under development. Because the systems were designed for general consumer use, it was possible to ask almost anyone (including the designers themselves) to act as a test user and receive relevant feedback.

Evolutionary systems development – the systems development should be both iterative and incremental. As mentioned above, the three individual systems were three iterations. For each of the systems the knowledge and technology developed in the previous iteration were used or re-used. A proper analysis of user requirements, needs, and use of the built technology was done between each system. However, within each system design no explicit iterations were made with specific requirements analysis, implementation, and evaluation phases.

Simple design representations – the design must be represented in such ways that it can be easily understood by the users and all other stakeholders. Due to the relative small scale nature of the project, no special emphasis was made on design representations. As mentioned above, no users were involved directly with the design, and there were no organizational or financial stakeholders to present the design to. However, because the work was done in an academic context, the designed systems and the user studies were presented as research results [I, II, III, IV, V], so in that context the system was well represented.

Prototyping – early and continuously, prototypes should be used to visualize and evaluate ideas and design solutions in cooperation with the end users. Low-fidelity prototypes were used in the design, as well as other means to support requirements elicitation. However, the low-fidelity prototypes were not evaluated by the users, which only evaluated the finished systems. Also, as the size of the system design groups was less than ten, there was no strong need to ensure a common understanding between the group members by prototypes.

Evaluate use in context – baselined usability goals and design criteria should control the development. No usability goals were explicitly documented in the design. However, the usability was the driving force of the development, especially in the later iterations/systems where the new technology was in less focus. Use of the systems was evaluated in context (see [II, IV]).

Explicit and conscious design activities – the development process should contain dedicated design activities. As mentioned above, the driving force behind the design of the systems was usability. The user interface and the user interaction were the primary drivers for the rest of the system, and practically all design activities were based on requirements for user interface and interaction.

A professional attitude – the development process should be performed by effective multidisciplinary teams. All teams involved in designing the system had a professional attitude towards the design. The teams were also multidisciplinary, including professional software engineers, academic researchers, usability experts, and graphical designers. The design teams had both industry professionals and

academics.

Usability champion - usability experts should be involved early and continuously throughout the development lifecycle. There was no dedicated usability expert on the teams. However, most members of the design teams had either an education in usability design or professional experience, or both. The concepts of usability or user-centered design were not alien to anyone involved in the design.

Holistic design – all aspects that influence the future use situations should be developed in parallel. As the designed systems were meant for general consumer use, there was little possibility to influence the future use situations, tasks or practices of the users. On the other hand, the situations of use were studied, mainly the typical hardware, software, and social environments in snapshot photography, and thus a holistic view on the use was adopted in the design.

Process customization – the UCSD process must be specified, adapted and/or implemented locally in each organization. The design teams had no documented in-house design processes to follow. However, the objective of this sub-chapter is to show that the principles of UCSD were followed to such a degree that the design can be called user-centric. Much of the principles were customized to fit the academic context of the projects as well as the small design teams and the non-commercial requirements of the early iterations.

A user-centered attitude should always be established. As mentioned earlier, usability and user-centeredness were the driving forces of the last two systems. Although the first system's goal was to use new technology, the user-centeredness was manifested in the user studies conducted on the built system. Also, the main objective of this dissertation is to look at metadata from the perspective of a snapshot photographer – a user of the systems.

To summarize, all of the principles of Gulliksen et al. were followed. Some were followed to the point (*e.g.*, explicit and conscious design activities, or user-centered attitude), and some were adapted to better fit the nature of the projects (*e.g.*, simple

design representations, or prototyping). The main tasks not followed were the direct involvement of users in the design process, and the explicit allocation of usability tasks or persons responsible for usability. Overall, the system design did follow the principles of user-centered design.

EVALUATION AND ANALYSIS METHODS

The principle of the methodological framework used here is to design and construct user-centric systems for photo sharing to be able to empirically study them in real life context, and to understand the impact of design decisions made. Therefore, the evaluation and analysis methods used in this research play an important part. Below is a summary of the user study and analysis methods. A more detailed description of the used analysis methods is in the original articles themselves.

The MMM-1 system's was given to fifty-five people to use for four months. Their use of the system was analyzed with three methods: a videotaped user interface test done on five test participants, a weekly survey on all participants for seven weeks, and two focus group discussions (altogether 15 subjects in the two groups).

The MobShare system was tested on two user groups for six weeks each (eighty-seven users altogether). Five members of each group were interviewed about their photography habits and social networks before, during and after the test period. The photographs, comments, and user logs were qualitatively analyzed after the user studies.

The PhotosToFriends was analyzed through anonym user data on roughly two thousand users. The data was used to study the social activity of a shared image gallery. Also, the differences between MobShare and PhotosToFriends were qualitatively analyzed from the perspective of commercial requirements on photo sharing systems.

SCOPE OF RESEARCH

The media studied in this work is mainly snapshot photography. The primary camera used in the user studies and in the systems is a camera phone. However, we argue that the results of this research apply to all snapshot photography independent of device. This work draws from several academic disciplines, such as cultural studies on photography and content-based image retrieval. However, this work lies mainly in multimedia and metadata research where we adopt a user-centric software engineering approach.

The technical implementation details of each system are left outside the scope of this research because the focus is on design issues rather than actual implementation. This is also true of the metadata and ontology discussed: the goal is to show general design issues rather than implementation dependent on metadata description language problems.

This work is engineering research. The intended audience for this work is software engineers and other designers of media sharing systems for consumer or snapshot use. This is to be kept in mind as we discuss the user trials and their findings: our goal is to use ethnographic methods to elicit new requirements for photo sharing systems. However, we believe that the findings do contribute to fields outside software engineering and system design.

CONTRIBUTION OF RESEARCH

The contribution of this work is the results of a user-centric approach into designing metadata for snapshot photography. This contribution has three parts. First, the understanding on social uses people have for their snapshot photographs: what did the test users photograph, with whom did they share the photos, and what kind of discourse emerged around the shared photos. Second, grounded on empirical evidence from the user trials, we describe how the different design decisions affect use: the use of galleries as the basic elements of sharing, the use of phone numbers in

enabling controlled sharing, the use of galleries and comments in facilitating discourse, and the importance of social awareness and notifications. Third, the implications that the first two contributions have for metadata for snapshot photography and to media metadata in general: the inherent problems in information associated with personal photographs, the strong coupling of metadata with its application, and the opportunity to use social activity to generate relevant information as metadata.

In the conclusions part of this dissertation we also discuss future issues and opportunities for future work in metadata for snapshot media in general: the requirement for a multidisciplinary and a multiple technology approach, the extra user burden of managing privacy, the opportunities presented by collaborative media authoring, the customer lock-in effect in not standardizing metadata for snapshot media, and finally as an answer to the research question, we present six dimensions that are essential in designing of metadata for snapshot media.

RELATED WORK

LITERATURE

This dissertation is a multidisciplinary approach to metadata for snapshot photographs. The related work used in this research is from studies on snapshot photography in the film era and current practices with digital cameras, from state-of-the art studies on the use of digital cameras and photos, studies on the state-of-the art in media and metadata technology, research done on mobile media, and commercial, as well as academic systems built for personal media management. As mentioned earlier, the methods used in this work are from software engineering and design-science research. This chapter positions this dissertation work in relation to other fields of study.

FILM, DIGITAL, AND MOBILE PHOTOGRAPHY

Snapshot photography has been studied in humanities and in sociology. There the focus has been on snapshot photography in history [16, 26], snapshot photography in society [9, 39, 78], family photos in communication [13, 64], and the role of technology in photography [27]. For a much more extensive review on studies in snapshot photography see, for example, Ulkuniemi [84].

The work listed above is critical in understanding the role of snapshot photography in our culture and how the changes in photo technology have influenced the photos and their use. However, the studies above are on film camera

technology and do not take into account the changes that digitalization of photographs has caused. More recent studies have looked at how people manage and share their *digital* photographs [28, 73], and how they are used, for example, in storytelling [4].

There are also studies that focus on mobile photography (*i.e.*, the camera is a mobile phone) and how does the fact that a mobile phone is always with and always connected change photography habits and the actual photos. An early study by Mäkelä et al. [56] looked at mobile images in communication in field trials. Another early study was made by Koskinen et al. [51] who gave mobile phones with attached cameras to users for several months. More recently there have been significantly more studies on mobile phone camera use (see, *e.g.*, [43, 49, 50]). The findings of these studies show that the "always with" nature of the mobile camera prompts people to take photographs in situations where they usually would not have a camera. Also, the possibility to share photos from the phone (*i.e.*, email or MMS, which are the picture sharing technology standard on current mobile phones) has enhanced the communicative nature of photographs.

These studies on people and photography have no direct connection to metadata. However, for the user-centric approach adopted in this dissertation work, these studies provide the background and empirical evidence for forming the requirements for metadata and associated systems.

MEDIA CONTENT METADATA AND CONTENT ANALYSIS

The literature on media metadata is, as mentioned in the previous chapter, focused on public or commercial media. Media metadata standards, such as MPEG-7 or Dublin Core, are not designed for media created by the user (*i.e.*, the consumer) for herself. Also, the photography specific metadata standards are either focused on public archival (*e.g.*, VALO ontology [85]) or for use by the professional or amateur photographer (*e.g.*, EXIF).

Another related field of research is content analysis of visual media (or content-

based image retrieval) (see, *e.g.*, [1, 77] for a review of the state-of-the-art). The objective of this field is to create technology for analyzing the contents of media by computation and in that way to extract information about the contents. The features extracted are, for example, color histograms, geometrical shapes, and salient points, or for video, such information as camera motion or shot boundaries. The technologies in content analysis of visual media are key in automating the generation of metadata by computation. However, the semantic gap is a critical point in the advancement of content-based retrieval [77]: the semantic level of the extracted information is much lower than the semantic level of user requirements.

An alternative to content-based analysis is the use of contextual information to generate the media content metadata. This is the objective of the MMM systems (version 1 in [I, 20], and version 2 [21]). Parallel to this approach, this dissertation proposes a user-centric approach: to study the uses people have for the (snapshot) photographs and bring that knowledge into designing and facilitating metadata creation. To the best of our knowledge, a user-centric approach has not been adopted in media metadata design or in content analysis of visual media.

SYSTEMS

Much of the work done on personal media management is through constructing systems. Commercial photo systems also play an important role in forming *de facto* standards and conventions in consumer use and in the state-of-the-art in personal media management technology. Here we reference to relevant photo management and sharing systems – both commercial and academic.

COMMERCIAL PHOTO MANAGEMENT AND SHARING SYSTEMS

The most popular and relevant personal media management systems that are PC products (in contrast to services on the Internet) are the ones referred to earlier, namely Adobe Photoshop Album and Adobe Photoshop Elements, Apple iPhoto,

Picasa by Google. The organization of photographs in the Adobe products is built around the concept of metadata (see Figure 3 on page 40). Apple iPhoto has a tagging functionality as one of the organizing functionalities and Google Picasa has the tagging also as an extra rather than a key feature. These services, due to their commercial nature, are designed to support snapshot photography with functionalities such as printing photos, emailing them, ordering paper prints, and re-using media in creations like slideshows, photo books, web pages, and photo collages. However, these PC applications, except for Picasa, provide little functionality for shared discourse or shared media spaces for the user. Therefore, these applications are primarily for individual use, the sharing is exporting the media away from the system by email, or by facilitating the creation of a public web page, or by "blogging".

In addition to the personal media management products, there are services in the Internet where people can upload their pictures and make them available for others. Often these services call themselves "photo sharing services". These services have the benefit over PC applications that they are accessible with regular web browsers and therefore require no new software, they provide a good alternative to sharing photographs by email (*i.e.*, no large email attachments and a user interface designed for viewing photographs), and they provide one virtual space or location for groups of people to interact and converse. Examples of photo sharing services are Flickr and Yahoo! Photos by Yahoo!, and Snapfish by Hewlett-Packard[11].

The last two systems built in this dissertation work, MobShare and PhotosToFriends, are similar photo sharing services as the ones mentioned here. Therefore, the research done in this dissertation contributes into designing photo sharing services, especially in designing private sharing, group discourse, and social awareness in the systems, as well as discussing the role of metadata. An example of

[11] http://photos.yahoo.com/, http://www.snapfish.com/

this is the Finnish photo sharing service Kuvaboxi[12], which is the current commercial version of the MobShare and PhotosToFriends services.

ACADEMIC PHOTO MANAGEMENT AND ORGANIZATION

The issues people have in managing their photo collection have been studied widely in academia by constructing prototype systems to facilitate the uses. The work can be divided into the following five approaches. These approaches are not exclusive, and some of them combine different methods.

The time and date metadata has been leveraged in organizing photo collections into clusters [18, 30, 34]. This approach takes advantage of the emergent phenomena in snapshot photography that people tend to capture photos in events, which have a clear beginning and an end (*e.g.*, parties, travels) and that people tend to take photographs in bursts rather than evenly distributed.

Location metadata is also used to organize personal photo collections [65, 66, 83]. Although current digital cameras do not have any location or positioning technology, mobile phones inherently have some knowledge of location due to the phone network. Also, external GPS devices can be attached to mobile phones to provide more exact location information. The location of the photograph seems to be an intuitive piece of information in organizing personal photo collections.

A third approach is to facilitate the annotation task for the user (*i.e.*, labeling or tagging photos). The user knows what information is relevant to a photo, and therefore, the semantic gap problem in computational metadata generation is avoided to a degree. On the other hand, the annotation of hundreds or thousands of photographs is a labor-intensive and error-prone task. Therefore, the objectives of the systems described in [47, 52, 76, 87] were to make the annotation task easy or automate it to a certain degree. For example, in the FotoFile system [52], face recognition was used to automate the annotation of people in photographs.

[12] http://www.kuvaboxi.fi/

Another approach is to facilitate the organization by advanced visualization methods and interface techniques. Bederson [5] uses quantum treemaps and bubblemaps, Kang and Shneiderman [47] use a concept they call "semantic regions", Kustanowitz and Shneiderman [53] apply bi-level radial quantum layouts, and Drucker et al. [24] combine labeling with visualization methods.

The technology developed in content-based image retrieval has also been successfully applied to personal photo management. For example, systems described in [31, 32, 52, 89, 90] take advantage of face recognition and detection in organizing photographs. Systems described in [31, 70, 72, 80] use low level visual features and image similarity measurements in organization.

The main organizing approach adopted in this dissertation work is *social*: the people who the photos were shared with. As it is discussed in more detail later on, the organization of photos in MobShare and PhotosToFriends is based on what photos were shared with whom. In addition, the time and date metadata in the photographs is used if available, and some basic visualization techniques are used to provide better usability.

MOBILE PHOTO SHARING

Practically all camera phones have in them the possibility to send picture messages – either by MMS or by email. However, messaging with pictures is often one-to-one communication, and most often from phone to phone. On the other hand, the inherent network connection and the "always with" characteristic of mobile phone cameras has made them a popular device for sharing photographs in the Internet. Publishing pictures directly from the camera phone into the Internet is often called "photo-blogging" or "mo-blogging". In addition to leveraging the direct network connection of mobile phones, academic research projects have taken advantage of the open programming interfaces in current smart phones. Unlike in regular digital cameras, it is possible to make applications where the camera function is only one feature among others.

The mobile photo sharing applications relevant to this dissertation are the mGroup [44, 45] system, which is a media sharing application designed for group use, and the MMM-2 [21] system (the version 2 of the MMM-1 system presented in [I, II]), ZoneTag [91] and the Merkitys-Meaning [61] systems, which are examples of leveraging contextual information in sharing media and generating media content metadata. The mGroup system is a client-server system, where users can create and subscribe into media stories that are made of text discussions and photos. It is designed for creating and sharing experiences within groups during large-scale events. MMM-2 is a photo sharing system where the user shares the captured photo immediately to other people either by email or posting to a website. MMM-2 facilitates the sharing by guessing the probable recipients based on the near-by Bluetooth devices at the time of capture and the user's previous sharing. Merkitys-Meaning is also a camera application, and it stores all available contextual information (location information, Bluetooth environment, calendar events, and user tags and descriptions) at the time of capture and lets the user upload the photo to the Flickr service. The contextual information is uploaded as Flickr tags. ZoneTag is a similar application that enables uploading a photo and its contextual information into Flickr. Both MMM-2 and Merkitys-Meaning use the same ContextPhone [71] platform in gathering the context at the time of capture.

POSITIONING THIS DISSERTATION

As mentioned in the beginning of this chapter, this dissertation work draws from several disciplines. Studies outside the field of technology on snapshot and mobile photography are used as the basis for understanding the role of photography in the lives of people who are not necessarily professional photographers or even call themselves hobbyists – the "snapshooters". This information is then applied to media content metadata research, where a user-centric approach is a novel contribution. Commercial online photo sharing is rapidly gaining popularity, and this work provides the designers of those systems empirical data on usage and design of similar systems and a critical perspective on the role of metadata. In the research field

of personal media management the contribution of this work is in organizing personal media based on social use (*i.e.*, selecting photos for sharing). And finally, to the research on mobile media sharing , and especially, leveraging context information in sharing or publishing the media, this work draws attention to the social uses and how contextual information should be seen to include the social context (*i.e.*, who were present or otherwise associated with the photo) at capture and in sharing.

MMM-1: CREATING MOBILE MEDIA METADATA

The Mobile Media Metadata system version 1 (MMM-1) was designed and built from January to September 2004. It was done in co-operation with the Garage Cinema Group at the School of Information Management and Systems (SIMS) at UC Berkeley, Futurice Ltd., and Helsinki Institute for Information Technology HIIT's research projects DE Core and Mobile Content Communities (MC2). MMM-1 was a photo metadata creation system for networked cameras that took advantage of contextual information and previously generated content metadata to facilitate the creation of content metadata immediately after capture. The system is described in detail in [I] and the user test findings in [II].

DESIGN

The objective of MMM-1 was to take advantage of the two special characteristics of camera phones: the inherent network connectivity and available context information. Also, to implement the system the open programming interfaces of the device were used. The research question that the system tackled was the media management problem and how it could be solved by having semantic content metadata stored with the photograph at the time of capture. However, it was identified that mere computational extraction of semantic content metadata from media does not provide high level information. And, on the other hand, the user has knowledge of semantic content information at the time of capture, but this information is effectively lost as time goes by. The same is true of the camera phone device: at the time of capture it has access to contextual information.

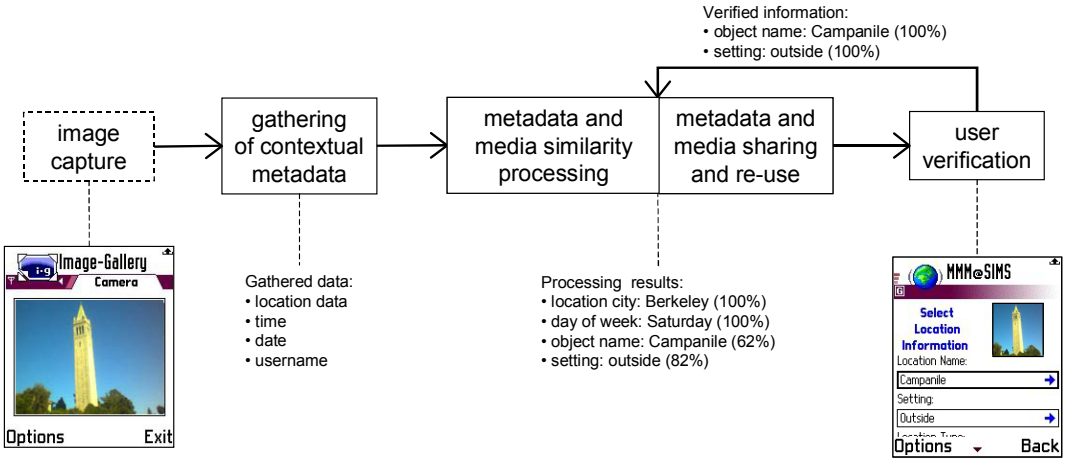

Figure 6. The metadata creation process of MMM-1. After image capture the client component of the system gathers available contextual metadata. Then that information is sent over the network to the server component that processes the context metadata and the media to come up with suggestions for user verification [I].

Rather than attempting to reconstruct semantic metadata by analyzing media after it has been captured, MMM-1 aimed to leverage contextual information and network effects to computationally assist users in creating useful semantic metadata at the point of capture. The system created, shared, and reused semantic metadata at the point of capture using the network connectivity of the camera phone. The content metadata was designed to be created by the following process which had the following principles (see also Figure 6):

1. Gather all automatically available contextual metadata at the time of capture.

2. Use metadata and media similarity processing algorithms to infer and generate new metadata for captured media.

3. Share and reuse media and metadata among users to facilitate metadata creation and new media applications.

4. Interact with the user during capture to verify and augment system supplied metadata.

At the end of the process the user would have the picture annotated with semantic metadata that was, to a certain degree, standard among all the users of the MMM-1 system.

IMPLEMENTATION

The system was implemented as a client-server model, where Series 60 compatible Nokia 3650 phones were the clients and an Apache Tomcat and HTTP servers the server components. To leverage the network effects of many people all the metadata and the captured media was transferred to the central server. In that way the freshly captured metadata and media could be compared to other people's metadata and media. For example, if 85% of photos taken at a certain location are of the same object, such as a tourist attraction, it is probable that the just captured photo in that same location is also of that object.

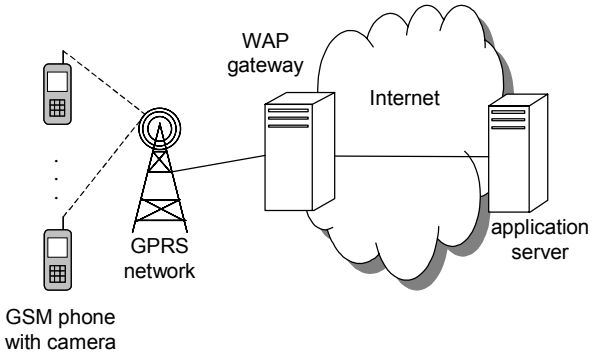

Figure 7. The physical architecture of the MMM-1 system [I].

Of the designed annotation process, the client program captured the time and date of the capture, the phone number of the phone that the photo was taken with, and the cell tower identification number. This was sent to the server that processed the cell tower information to provide a guess of locations, and another algorithm that processed the phone number to see who the people are that have been associated in previous pictures with the owner of the phone.

After the phone client had uploaded the information, it opened up an XHTML web browser to a web page given by the server that contained the server's guesses. From the web browser interface the user could verify or correct the guesses. See Figure 7 for the general architecture of MMM-1.

For describing the contents of the photos captured, the system used a simplified metadata structure (*i.e.*, an ontology). The structure was based on the faceted metadata hierarchy designed for the Media Streams system [19]. The structure had main categories, facets. The objective of these facets was to be as independent of each other as possible, in other words, one could be described without affecting the others. In the MMM-1 structure the facets were Person, Location, Object, and Activity. See Figure 8 for an example of the metadata structure used.

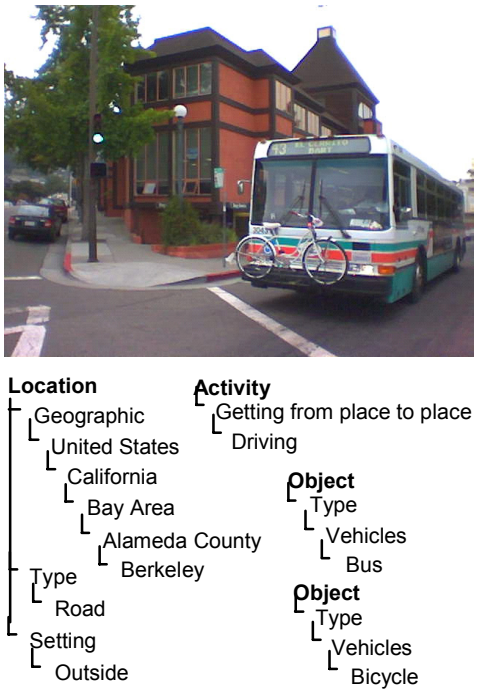

Figure 8. An example ideal image annotation in the MMM-1 system. The ontology used was hierarchical and faceted. Each description could include zero to infinity facets, which were Location, Activity, Person, and Object [1].

ANALYSIS

Once the system was constructed it was tested by deploying it to 55 users, who were graduate students and researchers at SIMS, for four months. All the test users were administered a weekly survey for seven weeks. In addition to this there was a user interface tests made with five participants, and two focus group discussions. A more detailed account of the system evaluation is in [I], and especially in [II].

RESULTS AND CONTRIBUTIONS

The main lesson learned from evaluating the system was that there has to be a purpose for the creation of the metadata so that the user perceives the benefits of the annotation effort. The MMM-1 system was initially deployed without any application for the metadata, and not until there was a desktop browser component built to browse the annotations, the users had no use or motivation to create the metadata. Although it sounds obvious in hindsight, it did emphasize that the effort people are willing to put into creating metadata is next to nothing, and the metadata has to have clear and understandable benefits, preferably immediate benefits, for users to go through any annotation task. In the systems that were constructed after MMM-1 (*i.e.*, MMM-2 [21] and MobShare [III]) this issue was tackled by concentrating on the user motives for sharing.

One of the main hindrances of the system was its reliability on the GPRS network connection. The annotation process was designed so that the user interaction worked over the network connection that turned out to be more unreliable and slower than the design anticipated. This in turn caused the user interaction to be so slow that it frustrated most users.

The so-called vocabulary problem [29] also became imminent. The users used existing words to describe different things, and vice versa. Also, the annotations turned out to be personal (*e.g.*, a stuffed toy monkey was annotated "George"), and therefore, not intuitive for anyone else.

It is worth mentioning, that in a more institutionalized metadata annotation scenario, such as creating metadata as part of a job in a public archive, the motives for doing the effort are simplified by the fact that the person is paid to do the work. In a consumer oriented scenario, such as the user tests of MMM-1, the motives to do the work have to come from the user herself. This also means that in consumer oriented scenarios for media metadata, the metadata has little value as such, or at least it is an overwhelming task to motivate people to create metadata just for the sake of metadata – the effort put into creating metadata has to be less than the benefits perceived from it. Therefore, the metadata has to be coupled with an application for it, and depending on the application, it affects and limits the kind of metadata that is created. For example, if the application for the metadata is in publishing pictures, then the motives for creating the metadata are related to the publishing task, and so is the created metadata as well.

This snapshot photographer perspective on metadata and the related metadata creation effort is the main contribution of the MMM-1 system to this dissertation. The system was designed based on the traditional library sciences principles for metadata: objective and structured information for a public audience. What the evaluation demonstrated was that in scenarios where the metadata is not used or created for the purposes of an organization, these principles do not hold true. As the user trials on the following MobShare system emphasized, the user-created information associated with photographs is contextual, personal, and dynamic.

Although the main lessons learned from MMM-1 tend to be lessons learned from design mistakes, one success was significant. The pictures taken with the phone showed a new photography behavior: the kind of pictures taken because the camera was always with. These were pictures were taken ad hoc and often in a humorous context to capture a special moment where they previously did not have a camera with them. The main objective of the following MobShare system was to study further the kind of pictures people take with camera phones and with whom do they share them.

MOBSHARE: SHARING PHOTOS WITHIN GROUPS

The MobShare system was designed and implemented from December 2003 to March 2004. The research done with it forms the main part of the research in this dissertation. The system was done in co-operation with Futurice Ltd. as part of the Helsinki Institute for Information Technology HIIT's research project Mobile Content Communities (MC2). MobShare was a photo sharing system designed for sharing mobile phone photos on a central web server. Its design principle was to facilitate the sharing of photos from mobile phones, and there was no special emphasis put on metadata. Also, the system was used as a commercial prototype for Futurice, and therefore, it was especially focused on end-user needs. The system is described in detail in [III] and [86]. There were two user trials done on the system. The findings of the first one are described in [IV] and the findings of the second one have not been published prior to this dissertation.

DESIGN

MobShare was built very much on the lessons learned with MMM-1 on end-user requirements for mobile photo sharing. Futurice, the partner in designing the system, had also experience with mobile photo blogging services and products in general. As mentioned above, one of the emergent findings with MMM-1 was the new photo taking behavior the users exhibited with the camera phones. Because the camera was always with, they tended to find new opportunities and situations to take pictures in. Also, these photos were used to relay light, transient, and often humorous messages to their friends and family [II]. Therefore, there was a strong need from the users to

share the images. This urge to share and communicate with photos was as noticeable as the need to create metadata was unnoticeable.

In addition to the previous work done on photo sharing, multimedia messaging (MMS) was used as a lesson learned. MMS had, at the time, enjoyed low commercial success in comparison to its high expectations. Because people seemed to have a strong urge to take photos and share them, it was an interesting questions why it was not used as much as expected. Our opinion of the low success was that MMS was designed to be a conceptual extension of text messaging rather than relying on metaphors and practices familiar in snapshot photography. This point of view was one of the main drivers in the MobShare design.

The architecture used in MMM-1 was also a clear influence in MobShare. To address the network problems encountered in the over the network user interaction, the MobShare phone client was designed to work so that when the network was used, it required no user attention. Secondly, the benefits of having a central server were taken into account in building the server side of MobShare: all pictures in one place, and desktop web browser access to photos, which meant a richer user experience and higher level of access than, for example, in phone-to-phone photo sharing (*i.e.*, MMS).

The main contribution of MMM-1 was the change of perspective for user-created media metadata, in other words, that the basic principles of the traditional library sciences metadata did not hold true in the context of snapshot photography. Therefore, the design of MobShare focused very little on metadata creation. However, the implications for metadata that MobShare and its user trials suggested were significant.

In addition to the lessons learned from MMM-1, a camera phone user study on ten Finnish users was conducted [69]. The user study pointed at the transient nature of mobile phone pictures in comparison to pictures taken with a traditional camera. It also emphasized the habit of sharing pictures in groups, in physical proximity, as well as by email. Thirdly, one of the findings was the personal nature of a mobile phone when compared to a camera. Traditional cameras are often shared property

within a family, but mobile phones are personal property. [69]

Finally, the literature on personal photo management and digital photography in general was studied to better understand the state of the art. Especially the user interface solutions in commercial systems and academic photo management applications for PCs were a fruitful source in designing MobShare.

IMPLEMENTATION

The system architecture was generally the same as in MMM-1 with the addition of a desktop web browser interface (see Figure 9). The client component was a Series 60 compatible application and the server side was implemented with J2EE and Apache Tomcat technology.

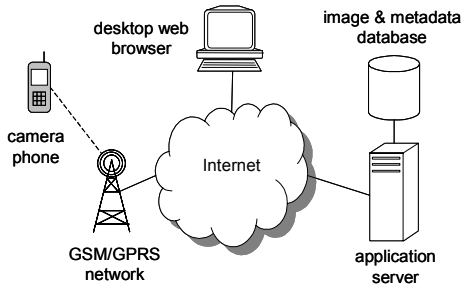

Figure 9. The general physical architecture of MobShare [III].

As mentioned above, the overall architecture was a client-server model. Initially the MobShare client program had the photo taking and sharing in one dedicated program. However, this was in practice redundant because the photos could be taken also with the phone's own camera program, and the benefits of having a camera program integrated with the sharing were not used by the users. On the other hand, to tackle the speed and latency problems of the GPRS network the picture upload process was designed to be user-initiated: the taking of pictures was a separate task from sharing the photos. The objective was to decouple the photo capture and the photo sharing, partly to follow the traditional film camera process where they are

decoupled, and partly to keep the capture and the picture upload tasks separate so that the slow network would not affect the photo capture. See Figure 10 for the picture sharing process on the MobShare phone client.

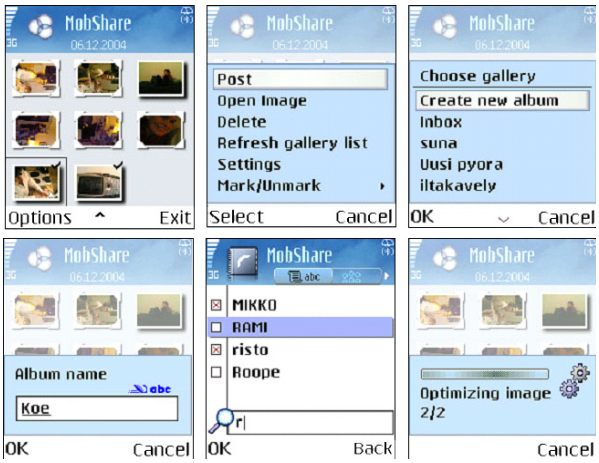

Figure 10. The sharing of pictures in the MobShare phone client. First the pictures are selected for posting, and then a new gallery/album is created for the selected pictures. The gallery is named and the recipients are selected. In the end the pictures are optimized and uploaded to the web server. [IV]

All the viewing, commenting, and organization of the captured and shared photos were implemented on the server side. This was because the server enabled desktop browser access (the users would not need a smartphone to view or receive pictures) and the desktop access enabled the use of a large screen and better input devices (*i.e.*, keyboard and mouse).

MobShare did not publish the images but limited access to galleries on an individual basis. For each gallery, the user had to explicitly choose the people who have access to the gallery, or in other words, who the gallery is shared with. There were no pre-defined lists of users (*i.e.*, buddy lists) that the user could take advantage of. Once a new gallery was created the people who the gallery was shared with got an SMS notification to their phone. The notification invited the recipient to visit the new gallery in the given URL, and it included the name of the person who shared the pictures and the name of the gallery.

To view the pictures, the user logged into the MobShare website. There she had her personal view of all of her own galleries and the galleries shared with her (see Figure 11). The galleries were organized in temporal order and according to ownership. The width of the gallery in the visualization was the time between the capture of the first and the last picture in the gallery. By selecting a gallery the system showed all the thumbnails of all the pictures in that gallery. By selecting a thumbnail the picture was shown in full size, and a textbox for comments. The pictures could be commented individually, and a summary of the comments for all pictures in the gallery were shown next to the thumbnails. There was also the possibility to comment the whole gallery rather than an individual picture. Galleries could also be created in the web browser interface of MobShare, where the functionalities were much more versatile (*e.g.*, adding and removing recipients of a gallery, creating galleries with no recipients, and deleting galleries). On top of the thumbnails there was a list of the people who the gallery was shared with. If the gallery was the user's own, there was also a list of all the people who had visited the gallery. Any pictures or comments that had been shared since the user's last visit were colored red.

Unlike public blogs or web pages MobShare was built for controlled sharing, meaning that the user has full control over who have access to the pictures. This also meant that for each gallery there was a distinguishable group of individuals associated with it.

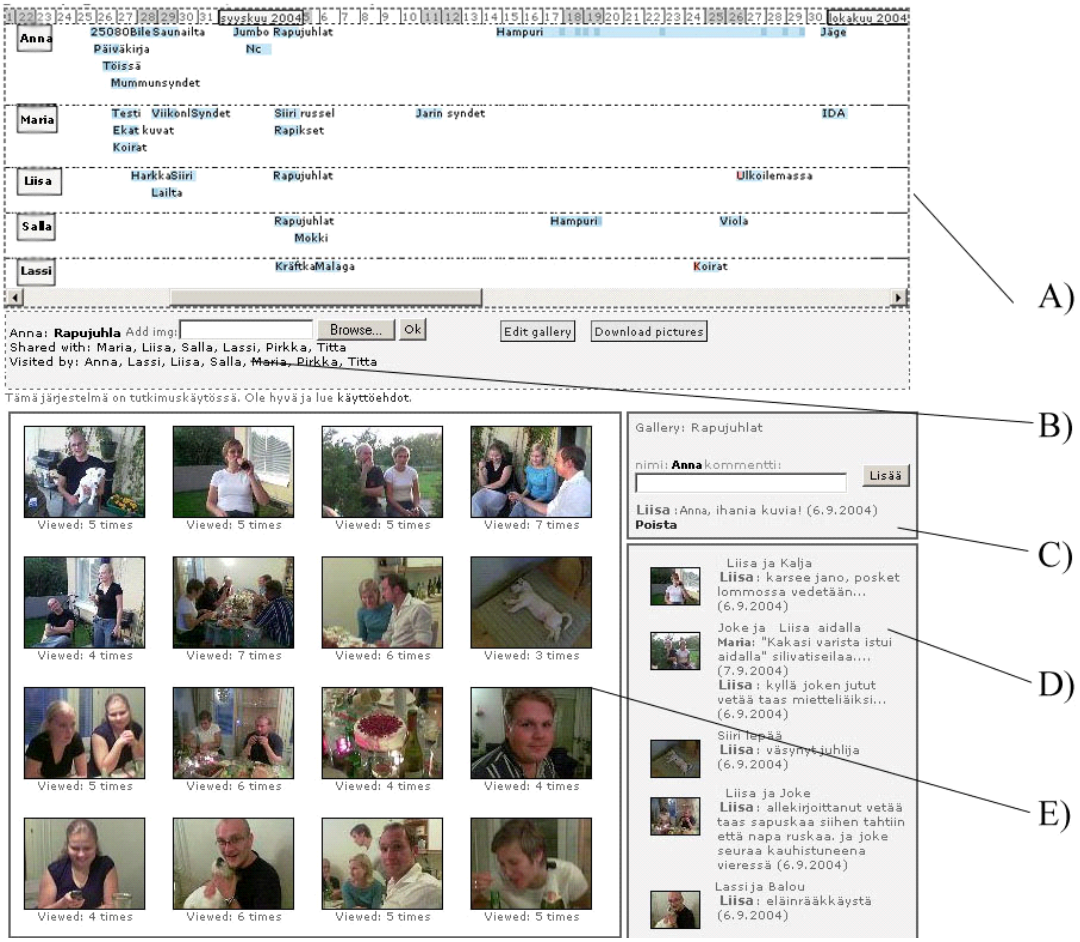

Figure 11. The web page interface of MobShare. A) User's own galleries, and galleries shared with her. B) List of people who the gallery is shared with and who have visited it. C) Gallery-level comments. D) Picture-level comments. E) The thumbnails of pictures in the gallery.[IV]

ANALYSIS

The main method of analyzing MobShare was two sets of user trials. For each trial there were five core users who were given Series 60 compatible camera phones (Nokia 6600 or Nokia 7610) and some guidance into using MobShare. It was emphasized that using the system was by no means compulsory. All of the data transfer costs for uploading the pictures were paid for, and all of the core users were paid a fee for volunteering in the study. Three interviews were arranged: before, in

the middle, and after the trial period. The interviews were about the users' photography habits and social networks, as well as their use of MobShare. In addition, the users were asked to fill out a diary, and the MobShare server logged the users' activity. Each trial lasted 5-6 weeks. The first set of users was a group of friends aged 25-26, four female and one male. The other group was a family where the core users (*i.e.*, the phones and fees were given to them) were the father and mother of the family, two children, and their grandfather. Both of the test groups shared photos also to people not hired as core test users (*e.g.*, friends and family). Altogether 87 people took part in the trials (10 core users and 77 secondary users). A more detailed account of the first set of user trials is in [IV].

RESULTS AND CONTRIBUTIONS

The main result on the MobShare user trials was a qualitative understanding of what kind of social uses people can have for mobile photos shared over the Internet, namely commenting and discussions, forming of groups to share photos with, the role of photo sharing as a means for keeping in touch, and the lifecycle of a mobile photo from capture to archival. It turned out that for the group of friends in their twenties the system provided a practical and entertaining way of communicating with friends and family. For the family, the system did not find such a need, because they saw each other daily and could share photos from the computer screen in the living room. An exception in the family setting was a daughter of the family who lived in a different city and sharing photos to her felt natural for them.

The pictures taken by the users were mostly of people. The galleries they shared were, in the group of friends, mostly of events (birthday parties or visits to friends), of travels to abroad, the photos of which were shared during the trip, and of themes (dogs or pictures taken at work). In the family setting the pictures were mainly of the children of the family, and the galleries a set of photos sent to someone in a gift-like fashion, mostly photos where the recipient was herself or had a special relation to the people in the photos. The children of the family created also galleries of artistic and experimental pictures. For the family, MobShare acted more like a convenient

technology to view and show the photos on the large screen of a PC. For the group of friends, MobShare was means for keeping in touch with their social network.

Whether it was over the network or at the living room, it was obvious that the purposes the users had for the photos were social. Even the picture taking was often social activity, for example, in the case where two children of the family took photos of a television screen while distorting the television picture by tuning the channels. The shared galleries also showed that the meaning of a picture is not necessarily created at the time of capture, but often during the discussions and comments during the sharing of the photos. Also, the meaning was contextual in the sense that the same picture could have a different comment, and therefore a different meaning, depending on with whom the picture was shared. From the perspective of snapshot photography this is not such a surprising finding. From the perspective of media content metadata it confirms that the information that people associate with their pictures is often social, and therefore, dynamic and contextual, as well as semantic.

One of the main contributions of the MobShare user trials, and in retrospect the MMM-1 evaluation, was an understanding on how design and implementation choices strongly influenced how the users could and would use the systems. Although, from a software engineering perspective it sounds obvious, in the context of photography it is not that intuitive. The difference between mobile phone photography and film or digital photography is the heterogeneity of available technology. The heterogeneity is much due to the nature of smart phones as multi-purpose devices. For example, the same device can be used to capture a photograph and send the picture as a message, not to mention all the other functionalities of phones that do not necessarily have anything to do with photography. In that sense digital cameras are more straight-forward devices with purposes and uses inherited from their film-based predecessors. Related to the multi-purpose nature of smart phones, a source of heterogeneity in camera phones are the open programming interfaces (APIs) that no digital cameras currently offer. The open APIs mean that potentially the camera phone can be used for anything, and examples of innovative new uses are games that use the camera as an input device (see, *e.g.*, [63]).

The heterogeneity of mobile phone systems from the point of view of snapshot photography becomes obvious when viewed through a generalized lifecycle of photos. To share photos taken with a mobile phone camera they can be shown from the screen or transferred to another device for viewing. There are currently four popular architectures for transferring photos from the phone: (1) to another phone over the network (*e.g.*, MMS), (2) to a PC (*i.e.*, the same procedure as with regular digital cameras), (3) to a network server over the network, and (4) to a printer over a cable or Bluetooth. Each of these ways of transferring pictures has unique characteristics. For example, transferring from a phone to another phone enables immediate sharing of pictures to practically anywhere the recipient happens to be with her phone. On the other hand, transferring pictures to one's own PC is often a familiar way of managing digital photos: once the photos are on the PC they can be edited, organized, published, etc. with the vast variety of applications. Also, there are no transfer costs between a phone and a PC or a printer, unlike often is the case in over-the-network transfers (*e.g.*, MMS or GPRS costs).

The lifecycle of a mobile picture can be divided into five subsequent phases. These phases are intuitive for any photographer and emphasize the heterogeneity of mobile picture sharing architectures.

1. **Capture** of picture using the mobile phone.

2. **Transfer** of pictures from the mobile phone.

3. **Sharing** of pictures, that is, making pictures available for other people to view and discuss, and as a recipient, being notified of pictures available for viewing and discussing.

4. **Viewing** the pictures involves not only looking, but also the related social interaction, such as talking about the pictures and commenting them.

5. **Archival** of pictures for later use, for example, a shoebox for paper photos, or a CD-ROM for digital pictures.

In Table 2 the most common mobile picture sharing architectures are compared in relation to the lifecycle described above. The lifecycle of a traditional film camera photos is included as a contrasting example.

Table 2. Photo lifecycles in common mobile photo sharing architectures. [IV]

Architecture	Capture	Transfer	Sharing	Viewing	Archival
MMS	Phone camera	Over the network when shared	Coupled with transfer. Shared individually	From phone screen	Phone's message in-box. No archival sup-port
Phone to PC (same as digital cameras)	Phone camera	Cable, memory card, or Bluetooth	Variety of sharing methods	From PC screen, via web browser, printed photos	PC's hard disk, web server, CD-ROM
Photo Blogging	Phone camera	Over the network when shared	Coupled with transfer. Shared on a web page	Via web browser.	Web page. Often no explicit archival support
Phone to Printer	Phone camera	Cable, memory card, or Bluetooth	Shared by showing printed photos	Tangible viewing of paper photos	Photo album, "shoe box", framed
Traditional Film Camera	Film camera	Film roll development	Shared by paper photos	Tangible viewing of paper photos	Photo album, "shoe box", framed

The table makes the following points:

- Lifecycle is technologically distributed over several devices.
- Coupling of lifecycle phases can be integrated in the system (*e.g.*, transfer and sharing in MMS are coupled into one function).
- None of the architectures are designed to have continuity over the whole lifecycle.
- Some transitions between lifecycle phases require user effort (*e.g.*, transferring pictures from phone to a PC).

Heterogeneity has been considered as an aspect of emerging ubiquitous and mobile computing environments [74], and to address the diversity of media and technologies in which people have interwoven current communication patterns [14]. In [IV], we examined the mobile photo lifecycle from a *systemic perspective* and took that perspective in our user study on MobShare. The study showed how the assignment of phases of the lifecycle to different platforms affected the social discourse and

activity around the photos. This opened a novel way to identify the interaction design qualities of heterogeneous systems as mobile photo architectures are: the *artful integration* of distributed functionalities assigned to different platforms, and *flexible and continuous lifecycle support* in the integration. We argue that coupling/decoupling strategies have a key role in implementing these qualities in inherently heterogeneous systems such as MobShare. To harness the best capabilities of mobile devices, and get over their known limitations, they need to be studied as parts in an ecology of devices.

PHOTOSTOFRIENDS: CONSUMER REALITIES AND SOCIAL ACTIVITY

The PhotosToFriends system (KuvatKavereille is the Finnish version) was commercial version of the MobShare system designed and implemented by Futurice Ltd. The objective of the system was to be a more commercially oriented and a polished version of MobShare, which was designed mainly as a research platform and a service prototype. Calling PhotosToFriends "MobShare version 2" would not be too far from the truth. The design and implementation work extended the work done on MobShare and the system was launched commercially in February 2005.

DESIGN

The goal in designing PhotosToFriends was to build a reliable system that would be easy to take into use, be used regularly (*i.e.*, similar to the regular use of email), be entertaining and generally a positive user experience. A crucial part in achieving this was to scale the system to handle thousands of users rather than hundreds as in MobShare, and therefore, to support all kinds of browser technology (MobShare supported only Internet Explorer web browser), use cases (MobShare did not support the ordering of print photos), camera technology (MobShare was built only for camera phone photography), and Internet notifications (MobShare notifications were sent only via SMS). Also, better general usability of the system was an important goal.

IMPLEMENTATION

The experiences and research done on MobShare was taken into account in both the design and implementation of PhotosToFriends. The architecture of PhotosToFriends was identical to MobShare and the biggest technical changes were done in scalability and reliability of the server component. The two biggest changes, from the perspective of this research, were the new gallery user interface, which replaced the horizontal timeline view of MobShare, and the support and facilitation of awareness and notifications of social activity.

The general usability of the system was emphasized by moving the location of the comments next to the pictures themselves, enabling notifications by email, showing in the gallery view the number of new pictures and comments for each gallery, and a new graphical layout for the whole system. See Figure 12 and compare it to Figure 11 to see the main differences between these two systems.

In addition to improvements to MobShare, new functionality was added, such as the possibility to order paper photos from the user's own photos but also from any photos shared with the user. The other major change was the support for uploading pictures from a PC with a web browser. On the other hand, some functionality was also discarded. Mainly the gallery level comments and the possibility to view several galleries as one or side by side. The reason for discarding these functionalities was that they were not used that much and the goal to minimize the complexity of the service for easy user acceptance.

Practically in all other aspects PhotosToFriends was identical to the MobShare system: the central architecture, the gallery-based organization of pictures, the association of people individually with a gallery, the controlled sharing in contrast to publishing, identification of users by phone numbers, and notifying users when another user shares a gallery with them.

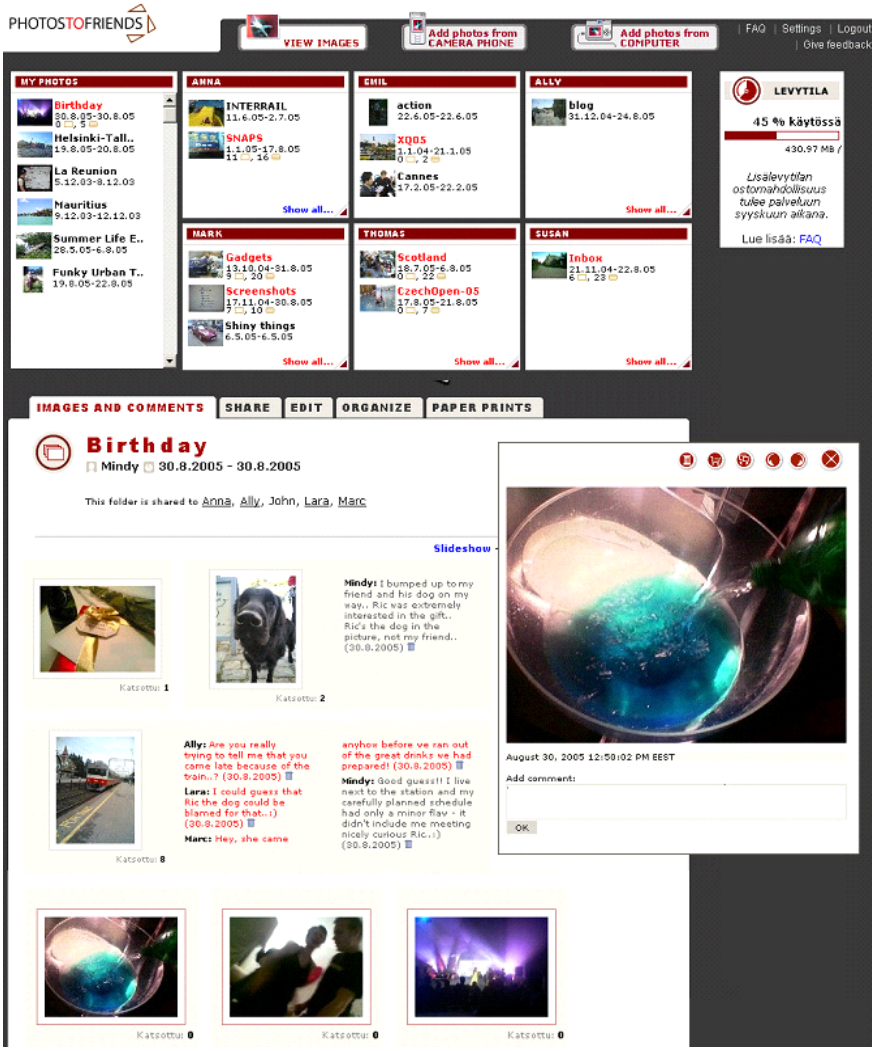

Figure 12 The web interface of PhotosToFriends. The user's own galleries and ones shared with her are shown on the top. New or unvisited galleries, comments, and photos are marked red. The people who have visited the gallery have their names underlined.[V]

ANALYSIS

The similarity of PhotosToFriends and MobShare makes it possible to use quantitative user data from PhotosToFriends to get a better understanding of the qualitative findings made with MobShare. The analysis of PhotosToFriends focuses on two goals: first, to qualitatively compare PhotosToFriends to the MobShare

system and discuss the effects and rationale behind the differences, and second, to leverage the quantitative user data of two thousand users to get a quantitative understanding of the phenomena identified with MobShare.

RESULTS AND CONTRIBUTIONS

The most visually notable difference between MobShare and PhotosToFriends is the changing of the horizontal timeline view (see Figure 11) into a user-based list of galleries (see Figure 12). The rationale behind the change is twofold: first, displaying the timeline view in MobShare used lot of computational resources and was not scalable as such, second, and more importantly in the light of this research, people do not have accurate time metadata in their digital photos. In PhotosToFriends it was possible to upload any digital images via the web browser, and quickly it became apparent that some users did not have the correct time and date in their photographs. Some users even had the time wrong by several years. Therefore, visualizing the organization of galleries based on the time and date of the capture was misleading and confusing when the time metadata was not accurate. One design choice would have been to use the time and date of the photo upload. However, this would have not supported the cases where the time was correct and the organization based on it very intuitive.

In MobShare, where the pictures were uploaded only from the mobile phone, the time metadata was often accurate, because people tend to keep the mobile phone's clock in correct time. This is most probably because people use the phone's clock as their watch and it also has other uses, such as, in call logs, which motivates the users to keep the clock in time. This change of the gallery organization echoes that, in practice, the time metadata of consumer photos can not be relied upon. Bulterman, in his article [11], argues the same.

Another design decision in PhotosToFriends also echoes the realities of consumer photography: the support for regular digital photos (*i.e.*, photos taken with digital cameras or scanned digital photos). Although currently there are more mobile

phone cameras sold in the world than digital cameras, it seems that people use digital cameras when they take pictures that they consider important. Another probable factor is that using the mobile phone in PhotosToFriends requires that the user has a smartphone and knows how to install client programs on it. Uploading digital cameras photos requires knowledge to transfer pictures from the camera to the PC and from the PC to the service via a web browser. To put it simply, transferring pictures from PC is much simpler, and people are more familiar with it than uploading from a smart phone. This heterogeneity of mobile photo systems is discussed in the previous section on MobShare in more detail.

However, the main contributions of the research done on PhotosToFriends are the design decisions and quantitative data on awareness of social activity. Unlike in public blogs or web pages the user in PhotosToFriends (as in MobShare) has full control over who have access to the pictures. As a result of this, for each gallery there is a distinguishable group of individuals associated with it (see Figure 12), and for each individual it is possible to record what pictures they viewed and what comments they wrote and when. This is possible to do for every user and user group because the service is a central web service rather than individual PC applications. In PhotosToFriends social activity was posting photos to the galleries, writing comments, clicking open galleries, and clicking open a thumbnail. Also, everyone could see who are the other people the gallery is shared with, and who has commented what and when. The owner of the gallery could see who has visited the gallery and when was the last time.

The awareness and notification of social activity grew to be a key feature in the service. The awareness and notification supported both immediate interest on the social activity, as well as long-time interest. In other words, immediately notifying the user that there is some ongoing social activity, and on the other hand, storing and showing the social activity of her photo sharing in a longer time scale. For example, the user could go back to old galleries and reminisce the comments and visits to photos from last year's birthday party.

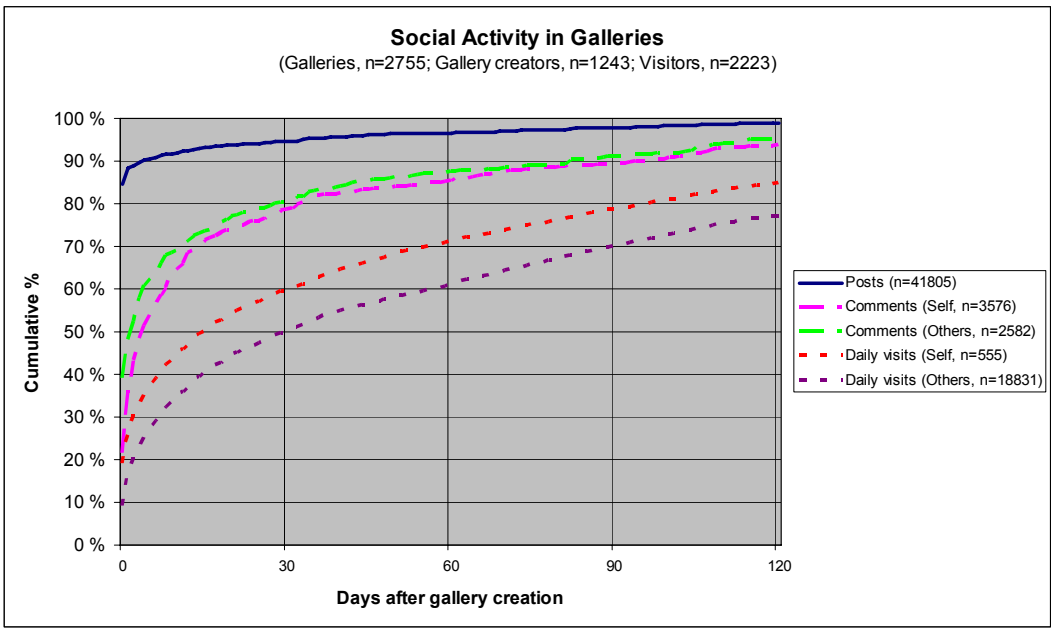

Figure 13. Cumulative data on social activity of 2755 galleries in PhotosToFriends from the day of the gallery creation to four months later. The data excludes galleries that were not shared to anyone. 'Self' in the figure means the owner of the gallery. The number of users in the sample is 2223. [V]

The awareness of ongoing social activity was done by enabling notification via text messages to users' phones and via email (*e.g.*, there is a new gallery, or there are new photos in a gallery). Once the user had logged in, the names of the galleries with new photos or comments were colored red, the number of new comments and pictures was shown in red, and inside each gallery the new comments and photos were marked red as well. The awareness of past social activity was supported by storing all comments and pictures, showing the latest visits to the owners of the galleries, and viewings of individual photos (see Figure 12). Figure 13 shows quantitative data how the social activity in a gallery develops over time. The figure shows both the immediate social activity in the very first days after a gallery has been created and shared (circa 90% of photo posts and 50-60% of comments in the first week). It also indicates the user interest in past social activity (the percentage of visits saturates towards 100% slower than posts or comments).

The quantitative user data indicates that a photo gallery in PhotosToFriends has

a relatively long lifetime before it is not visited any longer, and the reason for this is the social activity around the photos in the gallery (*i.e.*, the visits, the comments, and the number of views). The design decisions to facilitate the awareness of social activity support the view that the social activity around the photos is in itself interesting to the users, and that social discourse is a key characteristic when people share photos. The IRC-galleria and Flickr systems mentioned in the introduction section are other examples of social activity having an integral role in a picture sharing service.

RESULTS AND DISCUSSION

This chapter ties together the results presented above. The discussion is divided into three parts based on the research questions presented in the Introduction.

The first part discusses the uses people had for the mobile photographs in the constructed systems. These uses were mainly social in nature, and we discuss these uses by describing the photos taken, with whom the photos were shared, and what social activity there was around the photos.

The second part discusses how the design decisions made in the constructed systems affected the way the users took photos and shared them. The key technology discussed are galleries and temporal ordering of photos, controlling sharing by phone numbers, technology to support social discourse, and visualizing social awareness and sending notifications. We also draw attention to the heterogeneity inherent in photo sharing systems.

The third part goes through the implications the systems and their respective user tests have for metadata. The implications are the problems related to snapshot photography, the strong coupling of metadata and its application, and generating metadata as a side-product of social activity. We introduce the concept of social metadata as information in the form of metadata that supports the social uses identified. In the end we address the main research question, what kind of metadata would support the uses people have for snapshot photographs, by identifying six dimensions of metadata for snapshot photography.

In the fourth part of this chapter we discuss the validity, reliability, and

generalization of these results.

SOCIAL USES FOR MOBILE PHOTOGRAPHS

The systems built in this research were designed for communication rather than for storing or archiving photos for the user. Therefore, the uses people had for the photos in the systems were often social in nature, although not always, as in the case of the family, where the son of the family used MobShare for storing photos on the Internet. Here we describe and discuss the photos the users took in the user studies, the kind of galleries they created in MobShare and for what purposes. Also, the social activity around the photos in the MobShare and PhotosToFriends systems is discussed.

PHOTOS TAKEN

The contents of every photo taken in the MobShare user studies were analyzed (see Table 3). Majority of the photos had a person or several persons as the main subject (60%). This is consistent with Ulkuniemi, who lists the subjects in snapshot photographs to be family members, friends, relatives, and pets [84]. If the pictures of dogs are included in Table 2's "person" category, then 73% of the photos had a person, people, dog, or dogs as their main subject.

Table 3. The main subjects of the photos taken in the MobShare user studies. The categories are not exclusive: one photo can have a person and an object as its main subject. The purpose of the photo as described by the photographer and the photographer's relationship to the subject were taken into account in the categorization.

	Group 1 (friends) 589 photos	**Group 2 (family) 588 photos**	**Both groups 1177 photos**
Person or people	70%	51%	60%
Object	21%	24%	22%
Location, scenery	13%	8%	11%
Dog	8%	12%	12%
Artistic	0%	13%	11%

The second major category is "object". In the group of friends (Group 1) the objects were related mainly to food or drink (roughly half of the objects), and it was often the case that at social gatherings there was food or drinks involved and they became popular photography subjects. In the family group (Group 2), the objects were not dominated by food or drink. The objects in the family group were signs, posters, aesthetic objects, such as flowers, or people using objects, such as a friend playing a guitar, or a baby reading a children's book.

The location and scenery photos in Group 1 were mainly of travels abroad and from scenic places. In Group 2, the locations were aesthetic pictures of every day surroundings covered with snow, mountain sceneries photographed from a book, and sceneries from a family trip to a spa.

The artistic photos in Group 2 were either abstract photos, where the subject was not clear, or photos that the photographer himself described as artistic rather than any of the other categories. Most of the artistic photos were taken by the seventeen-year-old son of the family. His brother also took few similar photos imitating his older brother.

For both groups the photos taken were traditional in the sense that they mostly fit Ulkuniemi's descriptions: family, friends, relatives, pets, and travel. The artistic photos are clearly an exception to this. Also, the large number of objects photographed by both groups differs from the tradition. The reason might be that the combination of no additional costs for taking a picture and the ubiquity of the camera phone lower the threshold of taking photos of objects, especially the photos of objects where no person was interacting with the object (70% of photos of objects of both groups did not have a person as a main subject). In Group 1 the photos of food and drink were taken probably to document the event in addition by the people present. For example, at a birthday party where all five of the Group 1 test subjects were present, three of them took a photo of the birthday cake (see Figure 14).

Figure 14. The same object, the birthday cake, photographed by three people at the same event.

In discussing the contents of the photos taken, it is good to bear in mind that the photographers in both of these groups were using the MobShare system, although not all photos were shared in the system. This probably influenced the number and type of photos taken. Especially in Group 1, and in the birthday event displayed in Figure 11 and Figure 14, the users were actively using MobShare at the time. Therefore, they might have had the sharing in mind at the time of capture. In other words, the pictures were taken to reminisce the event in the future with the participants of that event, which is exactly what happened. Therefore, it is important to discuss also the sharing of the photos, that is, who did the users select as their audience.

SHARING OF PHOTOS

The main difference between the MobShare test groups was in the way they shared photos. Group 1 which consisted of five people (core users), who were friends before the tests, and 48 of their friends, relatives, and colleagues, used MobShare actively within the social network. Group 2, which was a family, used MobShare much less to keep in touch with their social network. They shared the photos, in addition to the other core users, to 29 of their friends, relatives, and colleagues.

Group 1 shared photos of events, themes, and travels, and also some photos as greetings to friends. All of the five core users had used email and text messaging to keep in touch with their friends and family, and MobShare replaced much of that

communication. Because MobShare was actively used to keep in touch with the users' social networks, the groups formed in sharing the galleries reflected their social relationships. Figure 15 and Figure 16 show the sharing graphs of two of the users in Group 1. The darker clusters reflect the emergent social networks: friends, family, and colleagues. For example, in Figure 15, the clusters reflect exactly the different groups of friends that the user had, and how these friends are treated as separate groups by the user in everyday life.

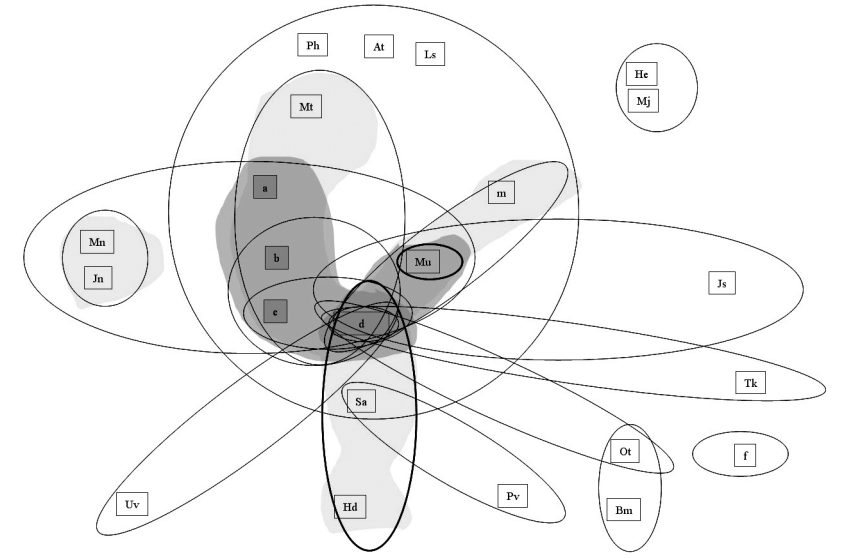

Figure 15. The groups created by a user in Group 1 by sharing photos in galleries. A circle denotes a gallery. The people who were grouped together twice have a lighter background. People grouped together more than twice have a darker background. The person in the center is the user's spouse. The cluster to the top right is the user's sister and mother, the cluster down is friends, the cluster from the center upwards is another group of friends, and the cluster on the left is a third group of friends.

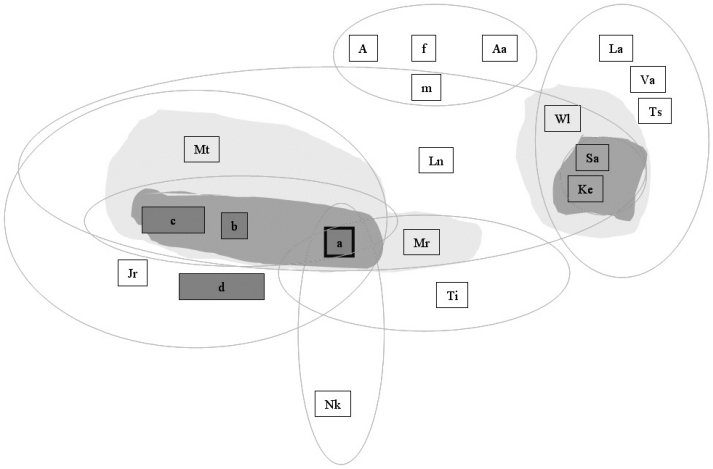

Figure 16.The groups created by another user in Group 1. The dark and light grey clusters on the left are friends, and the clusters on the right are colleagues.

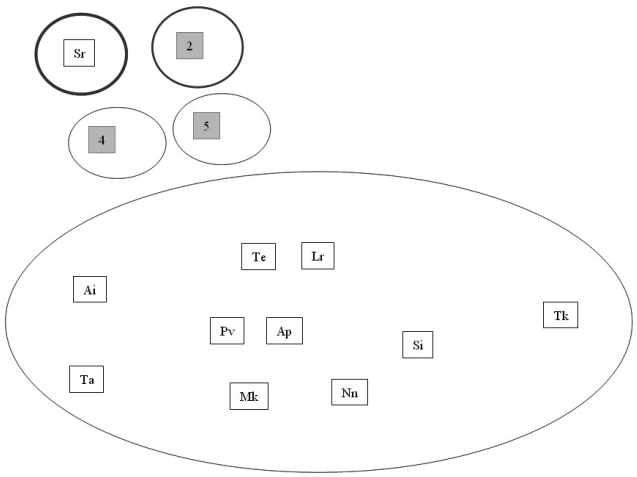

Figure 17. The groups created by a user in Group 2. The creation of groups did not show any clusters because no two or more people were shared together more than once.

In contrast to Group 1, the sharing in Group 2 did not form any clusters that reflected the users' social networks. Figure 17 and Figure 18 show two sharing graphs of Group 2. Because the initial setting for Group 2 was a family, the sharing

of photos happened in the family context and took place on the computer screen in the living room.

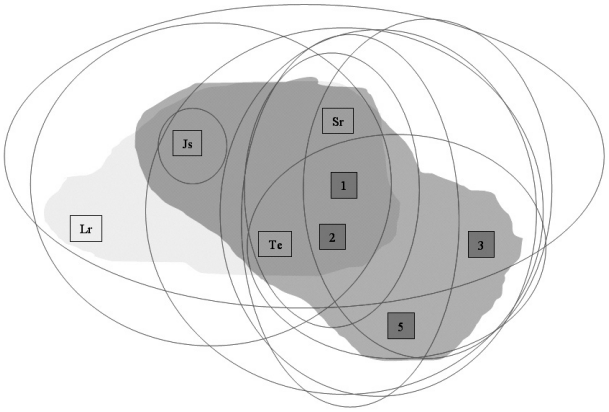

Figure 18. The groups created by another user in Group 2. In contrast to Figure 17, the user shared many galleries, but the people were always the same: siblings, parents, and grandparents.

Also, the galleries shared in Group 2 were not so often of events as in Group 1. The main type of gallery in Group 2 was a set of photos without any particular theme or narrative; only a set of pictures taken by the user. Mostly these galleries were named after the recipient. Some of the galleries were pictures of the recipient, and therefore, could be considered to be a form of gift giving. The seventeen-year-old son of the family, who took the "artistic" photos, shared galleries that were narratives of fictional events or art galleries of experimental photos. These galleries were shared to family members.

To summarize, the sharing in Group 1 reflected the way of keeping in touch familiar to the test users: because they met their different groups of friends (*e.g.*, friends from the university, old friends from high school, friends from work *etc.*), family, and relatives perhaps once in every two months. Prior to the tests, they had used email, text messaging, and calling as means of keeping in touch. During the

tests, MobShare was used for the same purpose. The photo galleries shared were mostly of events (parties, travels, happenings), and some galleries were sent as a set of pictures from different events in the form of a greeting.

In Group 2 the sharing in MobShare was not captured as much as in Group 1. The reason was that the users had no clear need to communicate with each other in the MobShare system because their social network was mostly their family, who lived in the same house or close-by. The oldest child of the family lived in another city, and she was shared photos by almost every user – to keep in touch. One user, the seventeen-year-old son of the family did use MobShare to create art galleries and showed them to the family members.

SOCIAL ACTIVITY AROUND PHOTOS

By social activity we mean the user actions tracked by the system that are shown directly or indirectly to the user or other users. In the case of MobShare and PhotosToFriends the social activity is the discussions around the photos, the red coloring informing the user of new photos or comments, the number of times a photo has been viewed, and the last time some recipient visited the user's own gallery. This activity has both immediate value and long-term value. The immediate value is showing what photo galleries are active now, and the long-term value is the storage of the activity for reminiscing purposes in the form of a group history.

The discourse in the MobShare system was designed to center around a gallery. Therefore, the discourse around the photos was within one gallery. As mentioned in [III], the types of discourse identified with Group 1 were storytelling, reports and self-documentation, greetings and thanks, and questions and opinion formations. In Group 2 there was less discourse, and the contents of it were mainly thanks to the sender for sharing the photos, inside comments that also told the sender that the recipient had visited the gallery, descriptions of artistic photos, and reports, such as mentioning that the little brother took this photo rather than the owner of the gallery.

For Group 1 the core user activity in the MobShare system was studied, also, by looking at the time and date metadata, the time of photos taken could be analyzed. It turned out that visiting the system was done almost regularly during weekdays. Within the weekdays the visits were done mostly in the morning and again in the afternoon. In contrast, majority (90.3%) of the photos were taken either in the evening (after working hours) or during weekends. In other words, the photos were taken outside work and the social activity around the shared photos happened during work. Much of this was because only two of the five core users in Group 1 had an Internet connection at home.

DESIGNING SYSTEMS FOR PHOTO SHARING

This chapter pulls together the main lessons learned in constructing the photo sharing systems. The objective is to provide insight on how certain design choices affect the use of the system.

SUPPORTING EVENTS BY GALLERIES AND TEMPORAL ORDER

A gallery of images (*i.e.*, a folder, an album) was the basic means of organization in the MobShare and PhotosToFriends systems. Individual pictures could not exist in the systems without a gallery. This design choice encouraged the sharing of several pictures at once rather than individual pictures. Fourteen percent of the galleries in the MobShare user tests had only one photo (17 galleries out of the 118 in Group 1 and 2).

The gallery creation design also played a key role in the number of pictures in a gallery and the type of galleries created. The camera and the picture upload were separate tasks, and therefore, the photo capture and sharing were also conceptually separate. In other words, the user could take photos without being prompted about sharing the photo – the sharing had to be initiated by starting the MobShare application on the phone. This supported the sharing of photos taken at events,

which had a clear beginning and an end (*e.g.*, parties, short travels, visits to friends), and often the users shared the photos of the event later in the evening or the next day. This design of decoupling the capture and the sharing is similar to the film camera sharing process where the time between capture and sharing could be several days or weeks. An alternative approach is to encourage the user to share the photo immediately after capture (see, *e.g.*, MMM-1 [I], MMM-2 [21], Merkitys-Meaning [61], or ZoneTag [91]). This approach is related to messaging (*e.g.*, text messaging) where people usually need to have more immediate interaction. The mGroup system [44, 45], for example, is designed for immediate messaging during an event, as opposed to the "post-event" design adopted in MobShare and PhotosToFriends.

Within the galleries in MobShare and PhotosToFriends, the pictures were organized in temporal order based on the capture time and date information, if available. This temporal order seemed to support best event-based photography where the captured photos have a documenting function of "this is what happened". For example, travel narrations are often documentations of the trip and changing the order of the events to contradict what really happened would be unconventional. However, no comparisons were made between temporal order and other means of organization in MobShare or PhotosToFriends.

On the other hand, the temporal order of photos somewhat limits the types of galleries possible. For example, to create a fictional story or an art gallery of images with particular order, either the timestamps of the photos have to be changed or the order of photo capture designed in advance. An example of an alternative approach is the IRC-galleria service where the photos are ordered according to the time and date of uploading the photo into the service. This approach supports displaying the activity and "freshness" of a photo gallery, because the latest activity can be presented first. This approach also solves the problem of relying on the users photos having the correct time and date. The finding we emphasize here is that the implemented organization method of a photo sharing system influences the kind of functions people have for the shared media.

DESIGNING CONTROLLED SHARING

The photos shared in MobShare and PhotosToFriends were not public. The user sharing the photos had to explicitly select the people with whom a gallery of photos was shared. The users were then identified by their mobile phone number. We call this controlled sharing. Using phone numbers had the benefit of connecting a trusted device into the system: the mobile phone and the personal SIM card (*i.e.*, the phone number). In contrast to email addresses, people often have only one or two phone numbers, each phone number is more secure, because if need be, the number can be tracked by authorities to an individual, and thirdly, having the phone number in the system enables messaging from the service to the user's phone. Because people most often have the phone numbers of their relatives, friends, and colleagues, this approach worked well when sharing within these social networks.

However, the controlled sharing did not suit the kind of photo sharing where publishing is important or the phone numbers of the recipients are not available or easily acquired. Publishing photos in the Internet has certain uses and practicalities. For example, publishing one's photos to create new social relationships can not be based on knowing the recipients beforehand (see the discussion on new phenomena in snapshot photography in the Introduction chapter). Also, people have situations where the group of recipients is known but their phone numbers are not, and often collecting them is too much work or socially awkward. Such situations can be, for example, parties where everyone does not know each other but would enjoy looking at photos afterwards, or larger groups of people at an event such as a conference. Often publishing on the web is then used as means of sharing the photos, and relying on emailing the web address to reach the intended audience. What is gained in easiness is then lost in privacy, and this is often managed by selecting photos suitable for publication either already at capture or before publishing.

SUPPORTING SOCIAL DISCOURSE

The text comments and discussions turned out to be a very important part of

photo sharing in MobShare and PhotosToFriends: the discussions kept the activity in a gallery alive for several days or weeks [IV, V]. Initially the MobShare system had three means of commenting: captions for each picture, comments for each picture, and comments for the whole gallery. However, it quickly turned out that these were redundant, and there was little difference in the contents of these different comment fields. Therefore, for simplicity, only the picture-level comments remained in later versions and in PhotosToFriends.

The other functionality that was important in the social discourse was commenting the pictures from the phone. In the MobShare user tests there was no functionality to add comments from the phone – all commenting had to be made through the web browser interface. This resulted in the users naming the gallery, which they were able to do on the phone, to be as descriptive as possible. In the user interviews, some of the users told this themselves. Also, at least in one occasion, a user withheld the sharing of the photos until she had access to a web browser to write comments on her photos. She told that without her comments the recipients would not have necessarily understood what the photos were about. In later versions of the PhotosToFriends phone client the commenting was made possible by building a special phone browser interface to the service. This also enabled the viewing of photos and comments from the phone.

VISUALIZING SOCIAL AWARENESS AND SENDING NOTIFICATIONS

Facilitating social awareness in the MobShare and the PhotosToFriends systems is one of the main design contributions of these systems. Constructing these systems and the associated user tests emphasized the requirement to visualize and communicate the social activity in the picture galleries. As mentioned above, the users in Group 1 regularly checked MobShare for activity, much like checking for new email. This was facilitated in PhotosToFriends by summarizing the latest activity by displaying the number of new photos and comments for each gallery in red color. Also, from the business point of view, the objective was to make people visit the

service regularly, and the social awareness had a key role in it.

Social awareness was extended also outside the system by notifications and special prototype applications. In the MobShare and PhotosToFriends system the users got a text message every time someone shared a gallery with them. This was changed into email messages in the current Kuvaboxi system due to cost reasons. Some prototypes were also designed and tested that displayed the social activity in MobShare, such as a mobile phone screensaver and an active wallpaper showing the latest photos and comments (see [IV, Figure5].)

HETEROGENEITY IN PHOTO SYSTEMS DOMAIN

Current systems, devices, and services that facilitate people's uses for their personal photos are very heterogeneous. In the basic lifecycle of a photograph there is a wide variety of technology in each phase: devices and programs for capturing the photo, means for transferring the photo to another device, applications and services to share or publish the photo, a myriad of ways to view the photo, and a variety of technology to archive the photo. To a designer of snapshot photo systems one of the main challenges is to address this heterogeneity.

The study on these three systems allowed us to define two qualities of interaction design in the heterogeneous consumer photo environments: *artful integration*, and *flexible and continuous lifecycle support*. By artful integration we mean combining existing systems and practices to new technology, such as making use of the existing phone book and practices associated with it, and leveraging the web browser to provide adequate temporal and UI resources for viewing and discussion, but also to enable more users as viewers and contributors. By flexible and continuous lifecycle support we mean designing the system both to support all phases of a snapshot's lifecycle and to do this in a flexible way to adapt to the heterogeneity. A key in implementing this are coupling and decoupling strategies. For example, decoupling the capturing and sharing of photos, coupling transfer and sharing, decoupling sharing and viewing, and coupling sharing, viewing and archival. The

coupling/decoupling strategies address the distribution of different tasks across applications on different platforms and devices, and we believe that this approach can be applied to other domains, in addition to online photo sharing.

IMPLICATIONS FOR METADATA

PROBLEMS WITH METADATA IN SNAPSHOT PHOTOGRAPHY

Metadata in snapshot photography faces certain problems, which are mostly due to the nature of the information stored and the motivations for the creator of the metadata to do the annotation. The information people would like to store with their photographs is highly semantic, dynamic, and contextual. The information involves people and their relationships, places and events important to people, meanings and shared experiences, generativity and family history, and as emphasized in the user trials of this work, socializing and communicating to others.

The problem arises in automating the metadata generation. Because the information is semantic, dynamic, and contextual, it is a very difficult to generate by mere computation. On the other hand, the user is quite familiar with the information she wants to store, and this information is probably understandable to her social network. Therefore, to generate this rich semantic information the user has to be involved in the process. This brings forth the issue of does the user perceive the benefits of having the metadata to cover the costs of investing time and effort in generating it. The tagging phenomenon seems to address this by not requiring strictly structured metadata and ontologies and by motivating the user to do the annotation by making it a tool for socializing and communicating.

The semantic, dynamic, and contextual nature of metadata in snapshot photography is not limited to the metadata nor to the time of media creation. The same problem is inherent also in the ontology. For example, if the ontology describes the relations between information such as people or locations, it is semantic,

dynamic, and contextual for the same reasons as the metadata. And, as the dynamic and contextual nature imply, the time of media creation is only one milestone in the lifecycle of the media, and hence, the metadata. People's relationships to other people, places, and events change over time, and the relationships and meanings can change depending on the context of use and the people constructing the meaning. This is nothing new in snapshot photographs: a single photo can have different meanings, feelings, and purposes in its lifecycle from capture, to viewing after few days, to reminiscing after twenty years.

Another significant problem with metadata in snapshot photography is in sharing the metadata. In addition to the problems mentioned above, sharing brings forth the vocabulary problem (*i.e.*, different people using different words to describe same things, and same words to describe different things). Even people with very similar backgrounds have the problem. For example, the same person can be "dad" for the son and "husband" for the wife.

The last problem listed here is that even the most basic automatically generated metadata can not be fully trusted. This problem is familiar to anyone owning a digital camera and traveling over time zones: to keep the clock in the camera at the correct time to have the time and date metadata accurate. In general, the user has a responsibility of keeping the sensors in personal devices accurate.

Overall, having rich metadata in snapshot photographs is very much at the responsibility of the user. Because the user is the best source for the information and the judge for the correctness of the information, the user has to be involved not only in the generation of the metadata, but also in the *management* of the metadata and the ontology. In other words, in relation to the user's own photo storage, the user has the responsibilities of a librarian.

COUPLING OF METADATA AND APPLICATION

In the snapshot photography domain, the application of the metadata defines what information is stored as metadata. This is because metadata as such has little

value to the user: the user can remember most of the information without explicit storage (*e.g.*, the interviewed user who did not write people's names into photo albums because she would not forget them) and writing down that information is therefore not motivated. To motivate the user to generate metadata there needs to be an application for it that demonstrates the benefits of annotation. Whatever the application is, it will select or restrict the kind of information generated. For example, if the application for metadata is making a photo book, then the user will generate the metadata that facilitates that task (*e.g.*, photos and the people, events, and places associated with them). If the application is for publishing pictures to form new relationships, then the user will generate metadata for that purpose (*e.g.*, construction of a public self-image, attracting people to the photos).

The user interface can also have an influence on the metadata created and maintained. For example, as discussed above, MobShare and PhotosToFriends organized the photographs in temporal order: the order of the pictures was based on the order of the time and date metadata. This implementation decision to couple the time and date information with organization in the user interface drew attention to the temporal metadata. The people using these systems to share their pictures were probably more aware of the temporal metadata in their photographs than people using systems that do not use or display the temporal metadata at all (*e.g.*, IRC-galleria.net). Also, the two systems mentioned did not provide any means for publishing the photographs. This design choice most probably affects the kind of discussions about the photos: less private matters are discussed or perhaps there is less discussion overall. This of course has an effect on the social information associated with the photographs: the comments and the people who have viewed the photograph.

The uses people have for their snapshot photos are very rich. They are different depending on gender, age, culture, personal history, situation in life, and so on. In addition, the technology in creating, organizing, sharing, and archiving photographs is heterogeneous. Therefore, the ideal of a common metadata ontology for snapshot photography such as the Dublin Core Metadata Element Set or MPEG-7, does not

seem possible.

METADATA AS A SIDE-PRODUCT OF SOCIAL USES

As discussed above, taking the metadata approach in snapshot photography has its problems. Information that would be useful and meaningful for the user can not be generated without some user participation, and the information generated is coupled to its application. Based on the research done on the systems presented in this dissertation we propose using the emergent socializing and communicating in photo sharing to generate metadata – social metadata.

Leveraging social activity to generate metadata is related to tagging. In tagging, such as in Flickr, the tags are social messages of a sort, which work also as means for organization. In the concept of social metadata we propose a more structured approach, which includes some rules and presuppositions about the information (see [V]). The idea behind both Flickr and social metadata is that social activity is motivated and it is mediated with technology.

Leveraging social activity addresses the motivation problem. It also indirectly addresses the contextual and dynamic nature of information: social information is inherently contextual and dynamic, it does not have the legacy burden of content information aiming to be objective and static.

Generating metadata through social media use is closely coupled with the application. For example, the gallery structures and means for discourse in MobShare and PhotosToFriends are integral in the kind of social information generated by users.

Generating metadata as a side-product of social uses is for future research, especially studying it in long-term user trials. Do people find social information relevant for storage, to what degree social information can be used in organizing media, and finally, what are the benefits of storing it as metadata? Within the scope of this work, these questions remain for future studies.

SIX DIMENSIONS IN DESIGNING SNAPSHOT METADATA

There are several ontologies and metadata standards for snapshot photography, such as the Dublin Core, EXIF, or IPTC, but none of these are designed with the snapshot photographer in mind. There are also several commercial media management applications that have tagging or metadata as means for organizing but they are product dependent and designed for personal use, not for sharing media. We address this limitation and answer the main research question of this dissertation by identifying six dimensions of metadata for *snapshot metadata*: privacy, ambiguity, dynamism, level of automation, level of structuring, and level of portability. Based on the findings of this work we see these six dimensions as the key issues that anyone designing metadata for snapshot photography (or personal media) has to address. These dimensions are important and relevant independent of the actual metadata. The first two dimensions (privacy and ambiguity) are related to the audience of the media and the metadata. The third dimension (dynamism) is a characteristic of the information itself: the meanings and uses associated with it are dynamic, and new information can be added as time goes by. The last three dimensions (automation, structuring, and portability) are about the creation process and the technical implementation of the metadata. Figure 19 is used as an example of a photograph with example metadata in discussing each of the dimensions.

The first dimension is the *privacy* of the metadata. At one extreme there is the private metadata that is meant only to one person, and at the other extreme is public metadata which is meant for anyone. In the middle of these two are different levels of privacy in relation to the audience or recipients of the media, such as, family, friends, or *ad hoc* groups like "everyone at the conference". The designed metadata should, therefore, address issues such as how the same picture can have different metadata depending on the audience. For example, if the photograph in Figure 19 is put onto a public web page, the user may want to keep the names of the people, comments, compilation information, and sharing information private. However, the user might want to publish the location information and would not mind if the time and date of the capture were public

Figure 19. An example photograph with example snapshot metadata.

The second dimension is the *ambiguity* in the information: how generally understandable the information is. For example, formal location coordinates are unambiguous and publicly understandable but a nickname associated to a location can be personal and ambiguous to a wide audience. In Figure 19 the user has given the name of the location as "Suvitupa", which is the name of the summer cottage few kilometers away. This is because for the user this area is associated with the summer cottage. On the other hand, the location coordinates and the country, city and place information inferred from the coordinates are publicly understandable and unambiguous. From the user's perspective the ambiguous and unambiguous information are complementary, and both have their uses: the coordinates can be used to associate location-based information to the photograph (*e.g.*, a map) and the user-generated name is for personal and intuitive organization. The ambiguity is, of course, critical in sharing the metadata: a more general and unambiguous description has better chances of communicating to other people. On the other hand, personal and ambiguous descriptions are intuitive and have rich meanings to people with shared experiences.

The third dimension is the *dynamisms* in the information. In other words, taking into account that the information associated with a photograph at the time of capture can change or accumulate over time. It is good to bear in mind, that photographs can have a lifetime of decades, even longer. For example, content and context information about a photograph often focus on the time of capture: people in the photograph, place of capture, location of capture *etc.* However, this can change over time as people's relationships to other people, places, and events in life change. For example, the user of the photograph in Figure 19 moves to live in the cottage in the picture. This would most probably change the user-generated location information in the metadata. Also, social information associated to a photograph is dynamic: with whom it was shared, what has been discussed about it, was it used in any compilations (*e.g.*, photo books or photo blogs), *etc.* In designing snapshot metadata, it should not be assumed that the information is always static.

The fourth dimension is the *level of automation* in the creation of the metadata. Is the metadata user-generated or automatically generated such as context data? For example, the location information in Figure 19 is user-generated and the coordinates are automatically generated. The metadata can also be created partly by the user and partly automatically. Nevertheless, the design challenge is to address the different characteristics of the metadata depending on the generation process. For example, metadata generated by the user can be ambiguous (*e.g.*, naming the location after the summer cottage kilometers away) and the user can even play with the presumptions and practices of the metadata (*e.g.*, annotating an object as "of desire" to communicate that the person would like to own the cottage in the picture). On the other hand, automatically generated metadata is not as semantically complex and this can be used as an advantage in designing uses for the metadata (*e.g.*, using the location coordinates to infer the country, city, and place information).

The fifth dimension is the *level of structuring* in the metadata. At its most basic form, the metadata can be free keywords with no hierarchy or any other structure (*e.g.*, the metadata "summer" and "old house" in Figure 19). A more structured metadata would include relationships and hierarchies between metadata elements, definitions of the type of information that is acceptable for a given metadata field,

and a formal ontology describing the metadata structure (*e.g.*, the metadata structure in Figure 19 could have the information that place – city – country are hierarchical, or that persons "Antti" and "Risto" are son and father). Designing the structure of the metadata is at the very basics of information organization: should the metadata be closer to the information retrieval school of thought or closer to the metadata and library sciences approach? The benefit of the non-structured approach is that it is flexible and lets the user define the contents of the metadata intuitively without any pre-defined rules to follow. The problem is that the free keywords with no structure are open to interpretation about the meaning of the metadata, and therefore, inferring new information is risky. The benefit of the structured approach is that there is less room for interpretation, and therefore, the metadata is more "effective" to use, and more concretely, the sharing and standardization of structured information is easier. The downside of the structured approach is that the structure assumes certain kind of uses, values, and types of information (*e.g.*, would there be rules for stepsons and stepfathers or would the structure assume a certain kind of family?), and therefore, is not as flexible as a free keyword approach. The designers of metadata have to address this tradeoff.

The sixth dimension is the *portability* of the metadata. By this we mean how application specific the metadata is. At one end there is standard universal metadata that can be interchanged between programs. At the other end there is application specific metadata that is not portable. For example, how would the information in Figure 19 be implemented into actual metadata? Examples of application specific metadata are the current versions of the commercial media management tools Adobe Photoshop Album and Apple iPhoto. Porting one's photo library and the metadata from one to the other is not supported[13]. This dimension has a direct connection to the business strategies of the products, as we discuss in the concluding part of this dissertation. One could say that the portability of digital images is currently very high

[13] Porting even within a media management application is not well supported: "As of 9/06 there doesn't seem to be ANY Apple supported way to join [iPhoto] Libraries. I think the partial support built into earlier versions of iPhoto may not work as well as it once did." (http://www.faughnan.com/digcam.html#LibraryMerge referenced 10.10.2006)

and successfully standardized from the snapshot photographer's perspective. A future question is will the metadata associated with images be as portable.

VALIDITY AND APPLICABILITY OF RESULTS

The main results and contributions of this work are the implications for metadata design in snapshot photography. The implications are based on the user trials made on the constructed systems. Here we discuss the external validity of the results presented and how they can be applied for commercial use and future academic studies. First of all, it is good to bear in mind that this work is engineering and design research and not, for example, social sciences research on people's photography behavior or the role of photography technology in everyday life. However, we sincerely hope that this work has a value outside engineering and systems design. The results presented in this work are meant to be input in designing technology for snapshot photography, especially online photo sharing and metadata. In the following, we discuss what issues the reader of this dissertation should take into account when generalizing these results and in applying them in his or her own work.

EXTERNAL VALIDITY AND GENERALIZATION

As agued earlier in this dissertation, designing metadata is a selection of all information available, and the design formalized some world view in the metadata and the associated ontology. The implications for snapshot metadata design in this work are based on the user trials presented and related literature on snapshot photography in general. Therefore, the implications are based on a certain view of snapshot photography in people's lives. A more extensive user trial would have given more insight into the state of snapshot photography in people's lives[14]. Also, the user

[14] After the work described here, we started a qualitative and quantitative study on snapshot photography in everyday life. The study was part of the *Immortalidad* project at the Helsinki Institute for Information Technology HIIT.

studies focused on Finland, and the cultural differences in photography and sharing media can have an influence on the metadata relevant for people. However, the related work referenced discusses several cultures, and the dissimilarities between the studies and the user studies in this work are not very significant. Especially as the photography technology used is the practically the same.

The first two systems constructed were designed for camera phones and could be used only by camera phones. The third system, PhotosToFriends, extended the support for any kind of digital pictures. However, we argue that the implications for metadata presented apply to all snapshot photography, independent of device. The results on designing photo sharing systems are not dependent on mobile phone technology. On the other hand, the types of pictures taken and the social activity in the MobShare user trials were influenced by the device, and this is taken into account in presenting the results of the user trials.

The implications for metadata presented have not been tested by construction and user trials. Therefore, we see important future work in implementing the implications for metadata in a concrete system and conducting user tests on it. Especially, the concept of social metadata should be further grounded on empirical studies to better understand whether it can be generalized outside the type of systems as MobShare and PhotosToFriends. Also, the design guidelines should be further iterated by designing an ontology for snapshot media and test that ontology in long term user trials.

APPLICABILITY OF RESULTS

We see that these results can be applied both commercially and in academic work. Because the research was done in close co-operation with a commercial photo sharing company Futurice, the commercial requirements for designing online photo sharing and metadata are present in the design decisions of MobShare and even more in PhotosToFriends. The most relevant design decisions in relation to commercializing the services were discussed in the chapter on PhotosToFriends. An

example of applying these results is the Kuvaboxi service, which is the current version of MobShare and PhotosToFriends. However, Kuvaboxi has not adopted a metadata approach in leveraging the social information generated. One issue in using standard content or social metadata in commercial photo sharing systems is customer lock-in, which is discussed more in the following chapter. In addition to designing metadata, we see the discussions on designing photo sharing and the results of the user trials directly applicable for commercial use: both in understanding how certain technology affects use and in understanding the end-users of these systems.

In academic work on metadata, we see the results as a novel perspective: in a domain such as snapshot photography the metadata design should be user-centric. In addition to this, the way people share, comment and discuss, and view photos (*i.e.*, what we call "social uses") can be used as information for organizing, re-using, and managing personal media. To the research on mobile media sharing, and especially, leveraging context information in sharing or publishing the media, this work draws attention to the social uses and how contextual information should be seen to include the social context at capture and in sharing.

CONCLUSIONS

Is there a future for metadata in snapshot photography? In the broad sense of metadata: yes, there is. It is already possible to automatically store relevant and useful information into images, such as location, time and date, and Bluetooth devices present. In the narrow sense of content describing metadata: not in the form of current metadata standards, which are currently practically unused by the regular snapshot photographer. For designers and researchers of current and future metadata standards for snapshot photography we draw the following conclusions based on this dissertation work.

A MULTIDISCIPLINARY APPROACH AND MULTIPLE ANGLES

Already the related work listed in this dissertation shows that relevant research is done in disciplines outside traditional multimedia and metadata. The user-centric approach presented here relies strongly in grounding the design in empirical data on actual use, and we believe that this perspective, which includes understanding users through cultural and sociological studies, will provide fresh insights into metadata design. On the other hand, the whole body of research in content analysis of visual media is integral in automating metadata generation and re-use. Also, more and more sophisticated media creation devices have been developed, some of which are already in consumer use. There are, for example, programmable mobile phones that make it possible to combine media creation with contextual information.

We also believe that the applications that use the metadata will have a key role in generating the metadata, and these applications should be useful and entertaining.

Also, as sharing is an important part of people's snapshot media use, designing for groups and communities should be taken into account.

Parallel to the multidisciplinary approach, we suggest a multiple angle approach to designing metadata for snapshot photography. By multiple angles we mean the different ways information is attached to snapshot media: user tagging, storing context information, content-based image analysis, and designing metadata and ontologies. All of these angles have some advantages and disadvantages. User-generated tags are well motivated, but not structured and hard to understand by computation. Storing context information is mostly automatic, but limited to the sensory information available. Content-based image analysis is also automatic, but it is limited to low-level semantics. A metadata and ontology approach is well structured and facilitates computational use, but in the domain of snapshot photography it is too labor-intensive for the user and the information is semantic, dynamic, and contextual. Therefore, we suggest that these angles are combined in designing metadata for snapshot photography, or snapshot media in general.

PRIVACY

The possibility to make one's personal photos public requires the owner of the photos to make decisions concerning privacy: personal privacy and the privacy of other people in the photographs. A snapshot photographer takes photos in non-public places where the ethics are dictated by social norms as well as law. Whether the photographer can make photos taken in non-public places public in the Internet is something the film photographer did not have to consider – not whether it is socially acceptable to publish or whether it is legal.

If, in the future, snapshot media has metadata associated with it in one form or another, it adds another layer of privacy management. In the situation where a person wants to make a photograph public, she needs to consider whether she will publish the metadata as well. The metadata can, for example, have information that is not implicit in the media, such as who were the people present but not in the picture.

This is currently possible in photos taken with applications such as Merkitys – Meaning [61].

Unlike media, which is designed to be viewed by a human, metadata is designed also for processing by computation. Therefore, a single photograph of a person in someone's public home web page does not necessarily offend the person's privacy. However, a single metadata description can be aggregated with other public metadata descriptions. For example, one person's phone's Bluetooth identification code in Flickr as a tag makes it possible to find and view all the pictures where the Bluetooth code is present (see [61]). This makes it possible to chart one person's location and company through time using Flickr tags.

The information associated with photos, whether it is context information or rich semantic metadata, has to be managed by the owner of the photograph. Automating the privacy management in metadata has the same problems as automating the generation of content metadata, and therefore, it requires user participation. This means that metadata in snapshot media adds another level of privacy management to photographs – another burden to the user.

COLLABORATIVE MEDIA AND METADATA

With more networked media creation devices, such as camera phones and cameras with WiFi-access, media creation becomes a shared task. As media can be immediately uploaded into a shared repository, it is possible to fade the individual authorship of media into a group authorship. The photos taken at a shared event become "our photos" rather than a collection of "everyone's photos". Rather than going into the ramifications of group authorship in snapshot photography culture, we point out the network effects in collaborate media and metadata. One prominent solution in lowering the metadata creation task is to distribute it to other people. This approach was in the MMM-1 system where other people's annotations could be re-used. However, it ran into the vocabulary problem, among other issues. The tagging in Flickr can be seen as a collaborative annotation task because anyone can annotate

any public photo. However, to what extent the work is collaborative rather than only the owners doing the annotation is an open question for future research.

Nevertheless, we see collaboration as one way of snapshot media creation, which has the opportunity to facilitate metadata creation by designing it into the shared creation and viewing task.

A METADATA STANDARD OR CUSTOMER LOCK-IN

The objective of media metadata standards such as MPEG-7, EXIF, and the Dublin Core is to standardize the way information associated to a media object is described. Once the format is standardized, new technology can be built to use media and metadata relying on the standard. Will there ever be a standard for snapshot media metadata? One problem is the personal nature of the media and the metadata. On the other hand, some information does seem to be universally useful for all snapshot photographers, such as, time and date of creation, location, name of the event, and the people in the photograph. The generation and management of both the metadata and the ontology are problematic, but we do not see that as a reason for not standardizing it.

Media sharing services and personal media management products are very popular and a growing business. People are becoming aware of the media management problem as their snapshot media accumulates, and sooner or later people need to stop capturing media or to start using an organization program. A similar trend is happening with online photo sharing services: more and more people are becoming aware of these services either by viewing someone's photos or by trying to find a more convenient alternative to sharing media by email. The more people use one system, such as a sharing service or an organizing program, the more locked-in they become. In addition to the media, all the socializing and communicating is stored in the system, and all the organization effort is done into this one personal media management software. Unless there is an easy way to migrate from one service to another without losing the social information or the organization

work, people are locked-in.

Storing the social metadata on media sharing, or any other metadata, in a proprietary format is a tempting way of keeping customers. It is possible to migrate one's photos from one service or program to another, but migrating the social information or the proprietary metadata structure is too much work. Therefore, in addition to the inherent problems in information about snapshot media, there might be little commercial incentive to make the metadata standard.

THE POTENTIAL OF SNAPSHOT METADATA

This dissertation began with a brief history of photo technology in snapshot photography. We argued that we are in the middle of a fourth major change in snapshot photography and this change is caused by digitalization of photographs and the availability network access. We pointed out three phenomena in snapshot photography that are the result of these technological changes: public audiences for snapshot photos, utilitarian snapshot photos, and tagging snapshot photos with free keywords. In the future, will metadata for snapshot photography be added as another major technological change contributing to the digital revolution in snapshot photography?

If all snapshot photographs, and other snapshot media, had content and social metadata stored with them it would change the use and role of media in our everyday lives. With metadata the media objects become versatile computable objects like any other digital object. However, the key question is what will be the information stored as metadata in the media. Currently, in the case of digital photographs, the information is what the EXIF standard stores and the benefits of that information are very limited. If the information would be more closely related to the uses people have for snapshot media, then it would amount to something significant. A personal media archive would become an archive of personal life: people, places, events, relationships, discussions, and social networks. This is what traditional photo albums already are, a personal or family history, but with the information as metadata the

archive becomes a computable database open to existing and future applications. For this future scenario to happen, there are open issues to address about metadata generation and management. Also, the question of the incentives for the commercial stakeholders (*i.e.*, camera device manufacturers, media management and editing application manufacturers, as well as media service providers) to work together to reach this is open. Currently, the information is application specific data inside products, services, and devices. For metadata to become a key technology in future snapshot media it needs to be as openly and freely movable and sharable as photographs are.

FIGURES AND TABLES

FIGURES

TABLES

REFERENCES

1. Aigrain, P, Zhang, H., Petkovic, D. 1999. Content-Based Representation and Retrieval of Visual Media: A State-of-the-art Review. *Multimedia Tools and Applications 3(3)*, pp.179-202. Springer, November 1999.

2. Alexa Internet, Inc. 2006. *Alexa Web Search.* http://www.alexa.com/ [cited 29.5.2006]

3. Baeza-Yates, R., Ribeiro-Neto, B. 1999. *Modern Information Retrieval*, ACM Press & Addison-Wesley.

4. Balabanović M., Chu L.L., Wolff, G.J. 2000. Storytelling with digital photographs. In the *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'00).* ACM Press, New York, NY, 2000, pp.564-571.

5. Bederson, B. B. 2001. PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. In *Proceedings of the 14th Annual ACM Symposium on User interface Software and Technology* (Orlando, Florida, November 11 - 14, 2001). UIST '01. ACM Press, New York, NY, pp. 71-80.

6. Bödker, S. 1998. Understanding Representation in Design. *Human-Computer Interaction, 13(2)*, pp. 107-125.

7. Boehm, B. 1988. A Spiral Model of Software Development and Enhancement. *Computer 21(5)*, pp 107-125.

8. Boivie, I., Åborg, C., Persson, J., Löfberg, M. 2003. Why Usability Gets Lost, or Usability In In-house Software Development. *Interacting with Computers, 15(4)*, pp. 623-639.

9. Bourdieu, P. 1990. *Photography, A Middle-brow Art*. Polity Press.

10. Bowker, G.C., Star, S.L. 2000. *Sorting Things Out – Classification and its Consequences*, The MIT Press.

11. Bulterman, D.C.A. 2004. Is It Time for a Moratorium on Metadata?, *IEEE Multimedia 11(4).*

12. Buur, J., Bödker, S. 2000. From Usability Lab to "Design Collaboratorium": re-framing usability practice. In *DIS'00: Proceedings of the conference on Designing Interactive Systems*, pp. 297-307. ACM Press.

13. Chalfen, R. 1987. *Snapshot Versions of Life*. Bowling Green, OH.

14. Chalmers, M. and Galani, A. 2004. Seamful interweaving: heterogeneity in the theory and design of interactive systems. In *Proceedings of the 2004 Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques* (Cambridge, MA, USA, August 01 - 04, 2004). DIS '04. ACM Press, New York, NY, pp. 243-252.

15. CNET news.com. 2005. Camera phone sales on upswing. http://news.com.com/Camera+phone+sales+on+upswing/2100-1041_3-5980049.html [cited 25.5.2006].

16. Coe, B., Gates, P. 1977. *The Snapshot Photograph – The Rise of Popular Photography 1888-1939*, Ash & Grant Ltd.

17. Cooper, A. 1999. *The Inmates are Running the Asylum: Why Hight Texh Products*

Drive Us Crazy and How To Restore The Sanity. Sams, Indianapolis, IN.

18. Cooper, M., Foote, J., Girgensohn, A., and Wilcox, L. 2003. Temporal event clustering for digital photo collections. In *Proceedings of the Eleventh ACM international Conference on Multimedia* (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, 364-373.

19. Davis, M. 1995. Media Streams: An Iconic Visual Language for Video Representation. *Readings in Human-Computer Interaction: Toward the Year 2000*, R.M.Baecker et al. (eds.), 2nd ed., Morgan Kaufmann, 1995; pp. 854-866.

20. Davis, M., King, S., Good, N., and Sarvas, R. 2004. From context to content: leveraging context to infer media metadata. In the *Proceedings of the 12th Annual ACM international Conference on Multimedia* (New York, NY, USA, October 10 - 16, 2004). MULTIMEDIA '04. ACM Press, New York, NY, pp. 188-195.

21. Davis, M., Van House, N., Towle, J., King, S., Ahern, S., Burgener, C., Perkel, D., Finn, M., Viswanathan, V., and Rothenberg, M. 2005. MMM2: Mobile Media Metadata for Media Sharing. In *Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI 2005)* in Portland, Oregon, ACM Press, pp. 1335-1338.

22. Dey, A.K. and Abowd, G.D. Towards a Better Understanding of Context and Context-Awareness, Presented at the CHI 2000 Workshop on the What, Who, Where, When, Why and How of Context-Awareness, April 1-6, 2000.

23. Dorai, C. and Venkatesh, S. 2001. Computational Media Aesthetics: Finding Meaning Beautiful. *IEEE Multimedia, 8 (4)*, October 2001, pp. 10-12.

24. Drucker, S. M., Wong, C., Roseway, A., Glenner, S., and De Mar, S. 2004. MediaBrowser: reclaiming the shoebox. In *Proceedings of the Working Conference on Advanced Visual interfaces* (Gallipoli, Italy, May 25 - 28, 2004). AVI '04. ACM

Press, New York, NY, pp. 433-436.

25. Foote, J., 2005. Kooks, Obsessives, Sturgeon's Law, and the Real Meaning of Search, *IEEE Multimedia 12(3)*, pp. 4-7.

26. Ford, C., Steinorth, K. (eds.). 1988. *You Press the Button We Do the Rest – The Birth of Snapshot Photography*, Dirk Nishen Publishing, 1988.

27. Freund, G. 1980. *Photography & Society*. David R. Godine, Boston.

28. Frohlich D, Kuchinsky A, Pering C, Don A, Ariss S. 2002. Requirements for photoware. In the *Proceedings of the 2002 ACM conference on Computer supported cooperative work (CSCW'02)*, ACM Press, New York, NY, 2002, pp.166-175.

29. Furnas, G.W., Landauer, T.K., Gomez, L.M. and Dumais, S.T. The Vocabulary Problem in Human-System Communication. *Communications of the ACM, 30 (11)*, pp. 964-971.

30. Gargi, U. 2003. Modeling and clustering of photo capture streams. In *Proceedings of the 5th ACM SIGMM international Workshop on Multimedia information Retrieval* (Berkeley, California, November 07 - 07, 2003). MIR '03. ACM Press, New York, NY, pp. 47-54.

31. Gargi, U. Deng, Y., Tretter, D.R. 2002. Managing and Searching Personal Photo Collections. HP Labs Technical Report HPL-2002-67, HP Laboratories, March 2002.

32. Girgensohn, A., Adcock, J., and Wilcox, L. 2004. Leveraging face recognition technology to find and organize photos. In *Proceedings of the 6th ACM SIGMM international Workshop on Multimedia information Retrieval* (New York, NY, USA, October 15 - 16, 2004). MIR '04. ACM Press, New York, NY, pp. 99-106.

33. Gould, J.D., Boies, S.J., Ukelson, J. 1997. How to Design usable sytems. In

Helander, M., Landauer, T.K., Prabhu, P. (eds.), *Handbook of Human-Computer Interaction*. Amsterdam: Elsevier Science B.V.

34. Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T. 2002. Time as essence for photo browsing through personal digital libraries. In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries* (Portland, Oregon, USA, July 14 - 18, 2002). JCDL '02. ACM Press, New York, NY, pp. 326-335.

35. Greenbaum, J., Kyng, M. 1991. *Design at Work: Cooperative Design of Computer Systems*. Hillside, NJ, Lawrence Erlbaum Associates.

36. Gulliksen, J., Göransson, B., Boivie, I., Persson, J., Blomkvist, S., and Cajander, Å. 2003. Key Principles for User-Centred Systems Design. *Behavior and Information Technology 22(6)*, 2003, pp. 397-409.

37. Hevner, A.R. 2004. Design Science in Information Systems Research, *MIS Quarterly 28(1)*, pp.75-105, March 2004.

38. Hillman, D. 2005. Using Dublin Core. Dublin Core Metadata Initiative, 2005. http://dublincore.org/documents/usageguide/ [cited 31.5.2006.]

39. Hirsch, M. 1997. *Family Frames: Photography, Narrative and Postmemory*. Harvard University Press.

40. *ID3 made easy*, http://www.id3.org/id3v1.html [cited 26.5.2006]

41. InfoTrends/CAP Ventures. 2005. World Wide Digital Camera Forecast.

42. ISO/IEC. 1999. *ISO/IEC 13407: Human-Centered Design Processes for Interactive Systems*. ISO/IEC 13407: 1999 (E).

43. Ito, M., Okabe, D., Matsuda, M. (eds.). 2005. *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life*. MIT Press.

44. Jacucci, G., Oulasvirta, A., Salovaara, A., and Sarvas, R. 2005. Supporting the shared experience of spectators through mobile group media. In *Proceedings of the 2005 international ACM SIGGROUP Conference on Supporting Group Work* (Sanibel Island, Florida, USA, November 06 - 09, 2005). GROUP '05. ACM Press, New York, NY, pp. 207-216.

45. Jacucci, G. and Salovaara, A. 2005. Mobile media sharing in large-scale events: beyond MMS. *Interactions 12(6)*, November 2005, pp. 32-34.

46. Jalovaara, V-M. 2006. Eihän niillä kännyköiden kameroilla tee yhtään mitään... vaiko sittenkin aina jotain. *MikroPC, 6/2006*. Talentum Oyj.

47. Kang, H. and Shneiderman, B. 2003. MediaFinder: an interface for dynamic personal media management with semantic regions. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM Press, New York, NY, pp. 764-765.

48. Kapor, M. 1990. Software Design Manifesto. In Winograd, T. (ed.), *Bringing Design To Software*. Addison-Wesley.

49. Kindberg, T., Spasojevic, M., Fleck, R., Sellen, A. 2005. The Ubiquitous Camera: An In-Depth Study of Camera Phone Use. *IEEE Pervasive Computing 4(2)*, pp. 42-50.

50. Koskinen, I., Kurvinen, E. 2005. Mobile Multimedia and Users: On the Domestication of Mobile Multimedia, *Telektronikk 3/4*, 2005.

51. Koskinen I, Kurvinen E, Lehtonen T-K. 2002. *Mobile Image*. IT Press, Helsinki.

52. Kuchinsky, A., Pering, C., Creech, M. L., Freeze, D., Serra, B., and Gwizdka, J. 1999. FotoFile: a consumer multimedia organization and retrieval system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the*

CHI Is the Limit (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999). CHI '99. ACM Press, New York, NY, pp.496-503.

53. Kustanowitz, J. and Shneiderman, B. 2005. Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts. In *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries* (Denver, CO, USA, June 07 - 11, 2005). JCDL '05. ACM Press, New York, NY, pp. 188-196.

54. Kyng, M. 1995. Making Representations Work. *Communications of the ACM, 38(9)*, pp. 46-55.

55. Laukkanen, T. 2005. *Modding Scenes – Introduction to User-Created Content in Computer Gaming*, Hypermedialaboratorion verkkojulkaisuja - Hypermedia Laboratory Net Series 9, Tampere University Press.

56. Mäkelä A, Giller V, Tscheligi M, Sefelin R. 2000. Joking, storytelling, artsharing, expressing affection: A field trial of how children and their social network communicate with digital images in leisure time. In the *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'00).* ACM Press, New York, NY, 2000, pp.548-555.

57. Manjunath, B.S., Salembier, P., Sikora, T. 2002. *Introduction to MPEG-7 Multimedia Content Description Interface*, John Wiley & Sons, Ltd.

58. Manovich, L. 2001. *The Language of New Media*, The MIT Press.

59. Mathiassen, L., Munk-Madsen, A. Formalizations in Systems Development. *Behaviour and Information Technology, 5(2),* pp. 145-155.

60. Mawston, N. 2003. Camera Phones Outsell Digital Still Cameras in H1 2003 and Beyond. In *Wireless Devices Strategies*, Strategy Analytics, Newton Centre,

MA, 2003.

61. Merkitys-Meaning, http://meaning.3xi.org/ [cited 25.5.2006]

62. *Mobile Imaging Report*, The Future Image, 2006. http://www.mobileimagingreport.com/Mobile_Imaging.htm [cited 25.5.2006].

63. Mosquitos, Ojom GmbH. http://www.ojom.com/?id=17 [cited 29.5.2006]

64. Musello, C. 1979. Family Photography. In Wagner, J (ed.), *Images of Information, Still Photography in the Social Sciences*, pp. 101-118. Sage, USA.

65. Naaman, M. 2005. *Leveraging Geo-referenced Digital Photographs*, Doctoral Dissertation, Stanford University, July 2005.

66. Naaman, M., Song, Y. J., Paepcke, A., and Garcia-Molina, H. 2004. Automatic organization for digital photographs with geographic coordinates. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries* (Tuscon, AZ, USA, June 07 - 11, 2004). JCDL '04. ACM Press, New York, NY, pp. 53-62.

67. Nielsen, J. 1993. *Usability Engineering*. Boston, MA: Academic Press Inc.

68. Norman, D.A. and Draper, S.W. 1986. *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, NJ.

69. Pesonen, J. 2005. *Näppäilykuva sai siivet - Kamerapuhelin valokuvan merkityksien muuttajana*, Master's thesis, University of Helsinki, Department of Education.

70. Platt, J.C., Czerwinski, M., and Field, B.A. 2002. PhotoTOC: Automatic clustering for browsing personal photographs. Microsoft Research Technical

Report MSR-TR-2002-17.

71. Raento, M., Oulasvirta, A., Petit, R., Toivonen, H. 2005. ContextPhone - A prototyping platform for context-aware mobile applications. *IEEE Pervasive Computing, 4 (2)*. 2005, pp.51-59.

72. Rodden, K., Basalaj, W., Sinclair, D., and Wood, K. 2001. Does organisation by similarity assist image browsing?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM Press, New York, NY, pp. 190-197.

73. Rodden, K., Wood, K.R. 2003. How Do People Manage Their Digital Photographs? In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'03)*. ACM Press, 2003, pp.409-416.

74. Russell, D. M., Streitz, N. A., and Winograd, T. 2005. Building disappearing computers. *Communications of the ACM 48(3)*, (Mar. 2005), pp. 42-48.

75. Schonfeld, E. 2005. Yahoo ups the ante, Business 2.0 http://money.cnn.com/2005/11/14/technology/flickr_biz20/ [cited 25.5.2006].

76. Shneiderman, B., and Kang, H. 2000. Direct annotation: A drag-amd-drop strategy for labeling photos. In *Proceedings of the Fourth International Conference on Information Visualization*, May 2000, p.88.

77. Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., and Jain, R. 2000. Content-Based Image Retrieval at the End of the Early Years, *IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12)*, 2000; pp. 1349-1380.

78. Spence J., Holland P. *Family snaps: The meanings of domestic photography*. Virago

Press, 1991.

79. Star, S.L. 1999. The Ethnography of Infrastructure, *American Behavioral Scientist 43(3)*, Sage Publications Inc, 1999. pp. 377-391.

80. Sun, Y., Zhang, H., Zhang, L., and Li, M. 2002. *MyPhotos*: a system for home photo management and processing. In *Proceedings of the Tenth ACM international Conference on Multimedia* (Juan-les-Pins, France, December 01 - 06, 2002). MULTIMEDIA '02. ACM Press, New York, NY, pp. 81-82.

81. Taylor. A.G. 1999. *The Organization of Information*, Libraries Unlimites Inc.

82. Tešić, J. 2005. Metadata Practices for Consumer Photos, *IEEE Multimedia 12(3)*, pp.86-92.

83. Toyama, K., Logan, R., and Roseway, A. 2003. Geographic location tags on digital images. In *Proceedings of the Eleventh ACM international Conference on Multimedia* (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, pp. 156-166.

84. Ulkuniemi, S. 2005. *Valotetut elämät, Perhevalokuvan lajityyppiä pohtivat tilateokset dialogissa katsojien kanssa*. Doctoral dissertation, University of Lapland.

85. VALO, eli Valokuva-alan ontologia, The Finnish Museum of Photography. http://www.fmp.fi/fmp_fi/muvieras/kirjasto/valo/index.htm [cited 31.5.2006]

86. Viikari, M. 2004. *Kamerapuhelimella otettujen kuvien jakaminen lähipiirin kesken*, Master's thesis, Helsinki University of Technology, Department of Computer Science and Engineering.

87. Wenyin, L., Sun, Y., and Zhang, H. 2000. MiAlbum - a system for home photo managemet using the semi-automatic image annotation approach. In

Proceedings of the Eighth ACM international Conference on Multimedia (Marina del Rey, California, United States). MULTIMEDIA '00. ACM Press, New York, NY, pp. 479-480

88. *World Internet Users and Population Stats*, Miniwatts Marketing Group, 2006. http://www.internetworldstats.com/stats.htm [cited 25.5.2006].

89. Zhang, L., Chen, L., Li, M., and Zhang, H. 2003. Automated annotation of human faces in family albums. In *Proceedings of the Eleventh ACM international Conference on Multimedia* (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, pp. 355-358.

90. Zhang, L., Hu, Y., Li, M., Ma, W., and Zhang, H. 2004. Efficient propagation for face annotation in family albums. In *Proceedings of the 12th Annual ACM international Conference on Multimedia* (New York, NY, USA, October 10 - 16, 2004). MULTIMEDIA '04. ACM Press, New York, NY, pp. 716-723.

91. ZoneTag Photos, Yahoo! Inc. http://zonetag.research.yahoo.com/zonetag/ [cited 2.6.2006.]