

# TECHNICAL RESEARCH REPORT

Motivating Annotation for Digital Photographs: Lowering  
Barriers While Raising Incentives (2004)

*by Jack Kustanowitz, Ben Shneiderman*

TR 2005-55



*ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.*

*ISR is a permanent institute of the University of Maryland, within the Glenn L. Martin Institute of Technology/A. James Clark School of Engineering. It is a National Science Foundation Engineering Research Center.*

**Web site <http://www.isr.umd.edu>**

# Motivating Annotation for Digital Photographs: Lowering Barriers While Raising Incentives

Jack Kustanowitz  
University of Maryland  
College Park, MD 20742  
+1 (301) 405-1000  
jkustan@umd.edu

Ben Shneiderman  
University of Maryland  
College Park, MD 20742  
+1 (301) 405-1000  
ben@cs.umd.edu

## ABSTRACT

Frameworks for understanding annotation requirements could guide improved strategies that would motivate more users to invest the necessary effort. We propose one framework for annotation techniques along with the strengths and weaknesses of each one, and a second framework for target user groups and their motivations. Several applications are described that provide useful and information-rich representations, but which require good annotations, in the hope of providing incentives for high quality annotation. We describe how annotations make possible four novel presentations of photo collections: (1) Birthday Collage to show growth of a child over several years, (2) FamiliarFace to show family trees of photos, (3) Kaleidoscope to show photos of related people in an appeal tableau, and (4) TripPics to show photos from a sequential story such as a vacation trip.

## Categories and Subject Descriptors

H.5 INFORMATION INTERFACES AND PRESENTATION, H.5.4  
Hypertext/Hypermedia

## General Terms

Design, Human Factors, Theory.

## Keywords

Annotation, Digital Photographs, User Interface, Human-Computer Interaction

## 1. INTRODUCTION

The bottom line: People can't be bothered to annotate photos. How many people are diligent enough or have enough time and energy to get a roll of film back from being developed, go over the pictures, and put them into albums, instead of just sticking the pictures in a shoebox somewhere? How many people go through their digital pictures and give each picture a unique file name in an appropriate directory instead of leaving them in the default directory created by the camera software?

Not many [9]. On the other hand, the two most useful features for coping with digital pictures – chronological sorting and displaying large numbers of thumbnails – are already available. And users' responses when asked about more advanced features reflect some interest, but indicate that they would not miss them if they were absent [22].

As a result, more and more people find themselves with thousands of digital photographs with little organization and little utility, and are resigned to gaining no more benefit or enjoyment from them than the thousands of printed photographs stored in shoeboxes around the house.

Annotation has the power to transform this semi-random collection of photos into a powerful, searchable, and rich record of people's lives and experiences [4]. Some of the opportunities that come from annotated pictures are described in the Related Work section below. But the main question remains: What motivates people to annotate photographs, and how can we as developers better entice people to spend valuable time adding meta-information to their photo collections?

Part of the problem is that users themselves do not understand what they can do differently once the pictures are annotated [9]. In a study of information capture, 10 modes of capture were discussed [3], and only about half of them could be done outside the digital realm. The existence of these exclusively digital modes provides good motivation for finding ways to take advantage of them in applications that utilize the unique capabilities of the digital medium.

This paper describes several techniques of annotation currently available, in order to build a good understanding of what is possible and what the strengths and weaknesses of some of the annotation technologies are. We describe a framework for annotation techniques, to better understand who does annotation and for what purposes. The goal is to lower barriers to annotations by making the interface easier, while increasing the incentives by enabling novel, automatically generated presentations. Finally, we focus on four applications that are made possible by well-annotated data, in the hope that they would provide motivation for users to spend time adding rich metadata to their digital photograph collections.

## 2. TYPES OF ANNOTATION

The goal of photo annotation is to create semantically meaningful labels and associate them with the photo. Captions (freestyle text like "Mom & Dad at their 25<sup>th</sup> anniversary") can be used for full-text searches and enable storytelling and dialogue-based sharing, but do not allow meaning to be inferred. Categorization (assigning a word from a finite vocabulary to a picture, like "sunset", or "Paris"), is better in the sense that searches are now possible on a finite set of words, but it still lacks the ability to bestow meaning (is Paris a place or the name of a person?). Without assigning independent semantic meaning to the labels, there is no way to ask for all people appearing in the photo collection, for example.

Annotation can also be in a non-text form, such as a sketch that abstracts the photo's contents for "looks like" searches where the input is drawn instead of typed [7]. For audio annotation, a melody line could be saved, to search an audio collection by humming into a microphone. For our purposes, though, we will focus on text, which is the predominant form and the most relevant for most people's interaction with photographs.

Annotation techniques can be grouped into three categories: Manual, Semi-Automated, and Automated. Each has advantages and

disadvantages, and we will discuss them in order of increasing automation

**Table 1. Annotation Techniques**

	Human Effort	Machine Assistance
<b>Manual</b>	Add structured annotations, with sufficient semantic information to be useful for retrieval.	Save annotations in a database.
<b>Semi-Automated</b>	Add freestyle annotations or captions. Potentially work with machine’s output in an iterative fashion.	Parse human-entered captions and extract semantic information.
<b>Automated</b>	Verify machine’s accuracy and make corrections as needed.	Add structured annotations using GPS, context, or recognition technology.

### 2.1 Manual Annotation

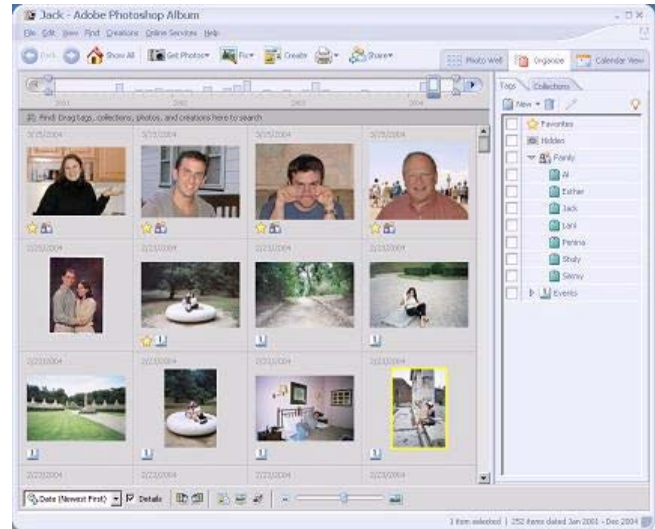
Most commercial software packages (Adobe Photoshop Album, ACDSee, Picasa, etc.) and web-based photo services (such as Snapfish and Ofoto) use manual annotation. They include the ability to set up a hierarchical list of categories (where depth of the hierarchy depends on the package), and add photos to those categories. They are an improvement over a regular file system in that a photo can exist in more than one place in the hierarchy. They also allow the addition of free-text captions, which while useful for online albums, can be difficult to search or to use for semantic extraction.

Manual annotation can provide the most accurate information, but it is also the most time intensive. The “group annotation” discussed below may mitigate the time requirement, but the user still needs to make one or more decisions for each picture.

Following are two examples of applications whose design for manual annotation lowers the barriers.

#### 2.1.1 Adobe PhotoShop Album

PhotoShop Album by Adobe is organized around a “Photo Well” into which photos are dropped when first imported. Users can define “Tags”, which are set up as a 2-level hierarchy, where either the tag name3 (“AI”, “Esther”) or the category (“Family”) can be dragged and dropped onto a picture in order to classify it.



**Figure 1**

PhotoShop Album also supports dragging and dropping in the other direction (picture onto tag/category), as well as multiple select drag and drop in either direction. A slider bar on the bottom of the screen allows fine control of the size of pictures in the photo well, allowing for selection and group annotation of even hundreds of pictures.

#### 2.1.2 PhotoFinder



**Figure 2. PhotoFinder**

PhotoFinder takes photo annotation one level deeper. Users can drag person names or other terms from a scrolling list and drop them onto a photo. Caption text can be placed on or under photos. PhotoFinder also associates annotations with coordinates within each photo [25]. This level of resolution potentially allows queries of “Bill next to John”, for example.

The downside is that the annotation process requires more attention to detail, since it is necessary to notice where on the photo the annotation is placed. Additionally, group-annotations cannot benefit

from coordinate information, since presumably the positioning of elements will be different in each picture in the group.

## 2.2 Semi-Automated Annotation

Semi-Automated Annotation starts with a manual process, presumably one that is easy and natural like voice annotation. It then goes through the manual annotations and extracts higher-quality, searchable metadata, which it then re-associates with the picture.

For example, SmartAlbum [29] assumes speech annotation of the photos has been done, and then proceeds to extract semantic information from the WAV file. It is limited by the accuracy of speech recognition, and will need to have some form of manual error correction, as the annotation algorithm uses heuristics to guess at event boundaries but is not guaranteed to succeed all the time.

Alternatively, the human-created caption can be parsed using a "Common Sense Knowledge Base" such as CYC [15] or OMCS [27], leveraging the implied context of a photo (i.e. recognizing a bride could cause an inference that the photo is of a wedding) to create a semantically meaningful annotation for future use. [17]

The MiAlbum system [32] uses a feedback mechanism to iteratively improve annotations as part of the search process. The first search returns random results, which the user grades for relevance, thus improving future searches. With each search and with extended use of the system, the quality of the annotations improves in an iterative fashion. [31]

## 2.3 Automated Annotation

Automated Annotation has the clear advantage that it happens with no user intervention, making it an attractive candidate. However, even in an ideal world, with perfect face recognition and shape detection, a computer will not be able to apply event labels, such as "Bill's 21<sup>st</sup> birthday party", or other heavily context-dependent annotations. Still, it behooves the designer to automate when possible, in order to minimize the amount of manual attention that is required.

The most basic type of automated annotation is done inside the digital camera, by applying the time stamp (which still requires that the user set the clock in the camera, no small feat for people with 12:00 blinking on their VCRs for 10 years). In fact, date and time have been shown to be the most important piece of metadata to record, as 92% of the subjects in a recent study had a specific time association with certain photographs [12].

An extension of this idea is to include a GPS receiver in the camera, and include a location stamp along with the time stamp. Microsoft distributes the WWMX TrackDownload application [35] that will use GPS data saved on a regular GPS device (and associated with a particular date/time) to stamp photos after they have already been downloaded to a PC. GPS receivers are attractive, but their inability to do location sensing in buildings requires some workaround, such as recording the last known location.

Another technology senses location based on the strengths of packets on an IEEE 802.11b wireless Ethernet network [14]. And finally there are proposals for getting geographic location by synchronization with cell phones that would be carried by the photographer. When the photographer uploads photos to a desktop or laptop that is internet connected, a link to the cell phone company database would produce time-synchronized location data.

The methods mentioned above use technologies with low susceptibility to errors, and thus with a high accuracy rate. Other approaches to automated annotation use methods that are more error-prone, but potentially yield more interesting data.

Some applications choose to use surrounding text as a way to generate extra metadata for photos [30]. Google, for example, uses this technique to automatically index vast numbers of images on the web in an automated fashion [11].

The Shoebox project [18] uses comparison of feature vectors (color, texture, location, and shape) to index images in an automated fashion. Advanced algorithms also exist for face recognition [37], and these can also be used to generate conjectures as to the contents of a photograph [36].

Aria [16] is a tool that links annotated photos to an email client. In addition to offering suggestions for relevant photos during the composition of an email, it is capable of adding annotations to photos in the collection as the email is written, based on keywords and information from a common sense database. [15][27]

As with machine transcription, the user can stop at that point, and decide to live with whatever the success rate might be, but it is usually worthwhile to take a second pass and fix errors or add additional information. In the best case, automated annotation saves most of the work and requires only touch-ups, but in the worst case it can require more time to "edit" the result than it might have been to just do manual annotation from the beginning.

## 3. A FRAMEWORK TO DESCRIBE MOTIVATIONS FOR ANNOTATION

Since photo annotation is used in widely varying contexts for very different purposes, it is helpful to categorize why different user groups perform annotation, with the hope that the understanding will help target applications that aim to make the task easier for a particular target group. In doing so, we use the division of target audiences into "Self, Friends & Family, Colleagues & Neighbors, and Citizens & Markets" [24]. We then discuss each of the categories in terms of what barriers they present to the creation of quality annotation, and what incentives can be offered to offset those barriers.

**Table 2. Motivations for Annotation**

<b>Audience</b>	<b>Description</b>	<b>Motivation</b>
<b>Self</b>	Annotation of photos located on unshared PC	Orderly personality, plan to use search & visualization tools
<b>Family and Friends</b>	Photos are shared, and family & friends benefit from annotation work	Recognition and appreciation, social value of sharing easy access to photographs
<b>Colleagues and Neighbors</b>	Local community projects	Improve quality of community, same as volunteering for other public projects
<b>Citizens and Markets</b>	Large payoff to work ratio, as one person's work is used by large numbers of people	Financial reward, or recognition and industry credibility if unpaid.

If annotation is being done for a single user, the barriers for annotation are greatest. Laziness or just not having time can cause annotation to be put off indefinitely, as other more pressing tasks take precedence. The benefits need to be clear, and it is for this user that many of the most creative applications have been created. Unfortunately, most people who create for the other audiences begin by creating for themselves, and barriers at that early stage of learning the technology can prevent users from progressing to the stage of creating for larger audiences.

Family and Friends add an additional motivation of an external audience; annotated pictures have more value in the larger group. Members of this group can more easily share, locate, and view photos that have shared emotional value for all of them.

On the community level, annotation increases the value of group photos, whether they are historical, biographical, or projective (plans for a new building or park area, for example). It might be accepted as a way of contributing to the group, comparable to volunteering to serve on a board of directors or planning committee.

Annotation for a worldwide audience holds the great promise of one person's work benefiting millions. The World Wide Web is the best example of this kind of technology, and there is currently much discussion on engines that will search that medium. For photos, free text search is not an option, and so the importance of annotation is correspondingly greater. Members of the world community can be motivated by profit (it may be someone's job to perform annotations), in which case the job of the software developer is to make them as productive as possible in their work, and make the process stress-free and enjoyable.

Additionally, since the target audience is so large, annotators gain widespread recognition and credibility by doing good work that can be used by many people, even if they are not receiving direct compensation for their work. The best model for this is the Open Source initiative [20], in which the individual gets public credit on a web site and can demonstrate publicly visible work as a reference when applying for jobs or consulting positions.

Individuals are the greatest challenge for the software designer, since without being strongly motivated to add annotations, they will favor more pressing projects. Therefore, especially in this case, the external motivation needs to be highest, and barriers need to be removed whenever possible. Specifically, the task of annotation needs to be:

1. Fun – since the task is being done voluntarily
2. Effective – i.e. result in a valuable product.

The absence of either or both of these ingredients relegates annotation to a “chore”, which will get put off as long as possible and ultimately not get done.

Very little has been written about the “Fun” component, although designers do strive for GUIs that are engaging and easy to use. Even though fun is an inherently subjective quantity, for a non-essential tool, it is imperative.

To be effective, annotation needs to facilitate the goals of photo sharing in general, which can be translated into the digital realm as: Remote Sharing, Sending, Archiving, and Co-Present Sharing [9]. It also needs to be powerful enough to enable various retrieval technologies, some of which are described below.

In designing annotation applications for the individual that 1) successfully remove barriers due to task difficulty, and 2) add motivations based on powerful and compelling applications, the designer encourages the individual to get started in a new area of technology, and increases the chances that the individual will proceed to create for larger audiences to the benefit of all.

#### **4. APPLICATIONS**

For many of the audiences mentioned above, motivation is critical, since there are no profit or job rewards associated with completing the task, and recognition comes not as a result of the annotation itself, but what the annotation enables. Following are several collages of photos, each organized slightly differently. What unifies them is the attribute of automatic generation – they are all views of potentially hundreds or thousands of photographs that can be automatically generated using existing annotations. The hope is that by creating applications that *assume* quality annotations, these compelling presentations would in turn motivate users to create the annotations.



Figure 3. FamiliarFace

#### 4.1 FamiliarFace

We wrote FamiliarFace as a prototype application to illustrate automatic generation of a pictographic family tree. The user chooses people from the list of people in their annotations, and defines relationships (parent, child, spouse, etc.) between them.

The program then goes to the larger picture collection (possibly numbering in the thousands), and retrieves thumbnails of all pictures of that person, and creates a collage with scrollbar for viewing them.

Finally, it pulls a picture from the pool that has been marked as "Favorite", and uses that as the primary picture (on top of each window, on white background, in the snapshot). Controls are provided for narrowing the focus either by generation (view 1, 2, 3, etc. generations), or by calendar (view from June 1999 – April 2000).

With minimal work (defining relationships), entire family trees spanning dozens of people and several generations can be built. With the ability to read GEDCOM [10] files (XML standard for genealogy), even the genealogical data itself can be imported, making the entire process automatic.

With any attempt to show large numbers of pictures on a single screen, space is an issue. The scroll bars shown above make an effort to allow for more photos to be shown; another solution involves

allowing zooming out to see an entire tree with hundreds of nodes, and zooming in to view a single photo [2].

A single user might use FamiliarFace to quickly find pictures of a certain person, and identify weaknesses in the collection (people with very few pictures taken of them) in order to focus more on them in the future. In anticipation of a family gathering, an elaborate collage could be generated spanning multiple generations and including hundreds of pictures filtered by quality.

A company org chart could be put together as well, treating the corporate hierarchy as a sort of family. This would be useful in letting employees get a look at colleagues in other settings, allowing for a stronger social and empathetic bond to be formed in the workplace. On the level of citizens and markets, a FamiliarFace type of interface could provide broad overviews of photo collections with some implied hierarchy, like decades of an artist's life or accepted periods of time in history (Renaissance / Classical / Modern, for example). Users could browse huge collections and choose photos for personal or professional use, possibly paying royalty fees once they find the right picture for their purposes.



## 4.2 Birthday Collage

Figure 4 is a manually-generated collage: It is a series of pictures of an infant taken at regular intervals, and assembled for sharing with friends and family. Just using the date stamp from the digital camera, such a collage could be automatically generated easily, using a random choice of picture from each month. With some additional “quality of photo” annotations, the best photos could be filtered out and a complete work generated with minimal user effort.



Figure 4. Birthday Collage

The birthday collage is a classic “friends and family” application, and it extends easily to the realm of colleagues and neighbors. For example, construction jobs, community gardens and flowers, and laboratory experiments are all “projects” that grow change at small intervals and whose arrival at larger intervals is marked. A good example at the citizens & markets level is the 100 year history of photos intermixed with Life Magazine covers to provide historical context, created by John David Miller at Intel.

## 4.3 Kaleidoscope

The Family Kaleidoscope is a prototype of an organized view of a set of photographs. It starts with a single individual in the center, and proceeds outward in concentric rectangles, each of which is related in some way to a previous level.



Figure 5. Kaleidoscope

A control panel on the left controls which picture(s) are selected on a given level. The model assumes indication of “Favorites” within

the annotations, to narrow down the choices from the complete set of photographs in the collection.

Beyond the Friends & Family application shown above (Figure 4), a similar program could include a “six degrees of separation” feature that would tell you at a glance who has appeared in photos with whom. For colleagues, conference highlight pictures/posters, showing connections between people and even between people and products or booths could be generated to summarize the conference and to market the following year’s event.

And on the market level, Hollywood collages of actors and which movies they’ve starred in together could be dynamically created for sale on fan web sites, or as an added feature on movie-oriented web pages.

## 4.4 TripPics

TripPics is a prototype of an application for displaying pictures that tell a sequential story. In this case, the pictures describe a trip through Italy, with major parts of the trip set off by separate boxes.

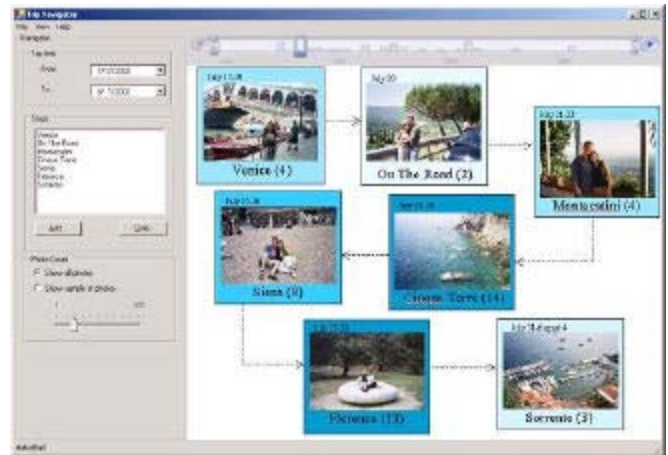


Figure 6. TripPics

In this screenshot, each “event” along the trip gets a single representative picture, automatically selected based on the “Favorite” tag, although the single picture could easily be replaced by the collage shown in FamiliarFace, to get a better idea of the entire collection of pictures from each part of the trip.

The pictures are shown on a white background, but they could also be superimposed on a semi-transparent map of the region traveled, with arrows or other indication of direction of travel.

A single user might use TripPics to quickly generate an album of a trip, which otherwise could take a lot of time. This album could be shared with friends and family, and provide more information than just a collection of pictures, even if free-text captions had been added. On a more global scale, different TripPic collages could be made public on a web site for people considering a similar trip, and could become integral parts of guidebooks to help evaluate destinations and sequences of travel.

## 5. RELATED WORK

Many other applications that require quality annotations have been proposed. The Personal Digital Historian (PDH) project [23] lays out photos on a round tabletop, and provides both an annotation engine and a novel display that takes advantage of the annotations to perform grouping, sorting, etc.

Fotonotes.net [8] is a web-based wiki-style interface that lets users create their own online collection of photos. Instead of annotating the photo as a whole, users create regions within the photo and attach story-style captions to each one. For example, in a picture of a living room, one caption could describe a painting and another a chandelier. Each region within the picture is then independently addressable, and can be viewed by URL as a cropped version of the original.

In the category of shared photos, Microsoft has an ambitious research project called WWMX, the WorldWide Media exchange [34]. It uses the GPS location stamp in the photo's metadata to associate a picture with a specific location, and then show a map of the world (zoomed to arbitrary detail) with pushpins of various sizes representing the concentration of pictures at that location.

An image browser that groups photos by visual similarity has been proposed [21], with mixed results. Some users appreciated the novel grouping (that would group pictures of sunrises, for example), but others were disoriented because the similar pictures seem to merge, making it harder to choose a specific picture.

When considering annotations done by a large group of people, questions of privacy, trust, and malicious users need to be addressed. In a controlled audience (a family or community), these issues are more manageable [15], as in the PhotoHistory of SIGCHI [28], at which annotation of a 20-year photo history was made public, with no malicious or inappropriate annotation. Even so, there were several requests to remove annotations and pictures due to privacy concerns.

On a larger scale, Wikipedia [33] is a large (6000 contributors working on 600,000 articles in 50 languages as of this writing) public annotation project, and does not seem to suffer from a large degree of malicious intent.

There are also several online review clearinghouses (Amazon.com book reviews [1], Epinions [6], etc.) where users can make annotations that are visible to the world. In these systems, some rules are explicit and are controlled by a moderator, and others are implicit and will result in peer criticism or reduced peer support.

Most of the online photo development companies (Snapfish [28], Ofoto [19], etc.) let the photo's owner to add captions. These captions are viewable (although often not searchable) by whoever the owner invites to share the pictures.

Ebay [5] is an example of a peer-to-peer financial transaction system, in which trust is critical. They have implemented a peer-review system, in which every seller has a rating that is a function of previous transactions. Sellers are highly motivated to provide quality goods and services, lest they be branded with a negative rating, which could affect future sales. While sellers with bad ratings can create new identities and "clear their name", they will also lose any positive rating as well as the sales history that shows their experience and reliability. This kind of system could be adapted to a large repository of digital photos to ensure quality and mitigate the potential for malicious use.

## 6. FUTURE WORK

A major challenge facing designers of digital photo consumer products is how to lower the barriers for manual annotation. When would a user prefer to annotate? While the photo is being taken? At download time? At "share" time? While the user is assembling an album? Each has its advantages and disadvantages: When users are downloading the pictures they are engaged in a basic annotation

exercise (choosing a file location), so that might be a good time. On the other hand, during sharing, users are communicating about the photos, and it seems reasonable to record descriptions then, for later semi-automated annotation.

Automated annotation depends on some advances in feature recognition, but a first step might be to recognize categories of objects (animals vs. people vs. scenery), and get that level of recognition into a consumer product. Perhaps even a limited automated annotation would spark an understanding of what full recognition might provide, and cause more interest in the process. The "cool" factor (wow, it knows which pictures are of my cats!) would also motivate people to put more work into annotations.

This paper discusses annotation of digital photographs, but photos are just a subset of the world of media, including voice, music, video, etc. These ideas could be extended beyond photographs, and the models reworked to cover a broader spectrum of media.

Some of the shared spaces discussed above in Related Work involve large numbers of people, with little malicious intent. Open source software and other communities form with widespread member involvement, and for the most part, people don't try to attack the system or damage it. At the same time, Microsoft frequently publishes security releases to prevent anticipated attacks against Windows, and email viruses have become ubiquitous. What makes a given system an appealing target for attack, and what is it about a community that encourages users to be helpful, protective, and to contribute their best? Why do some systems inspire malicious computer attacks with such as privacy intrusions and security breaches, and others inspire altruistic community-oriented sharing and mutual protection?

Finally, research should be done on what constitutes a "fun" interface, which might vary among applications and cultures. For individual annotation / individual use, it needs to have a visceral appeal that will make users excited to perform an otherwise optional task, and satisfied with quick and meaningful results.

## 7. ACKNOWLEDGMENTS

The authors acknowledge Adobe Corporation for funding research that contributed to this work. They would also like to thank Greg Elin for his feedback and comments.

## 8. REFERENCES AND CITATIONS

- [1] Amazon.com, <http://www.amazon.com>, Accessed on April 20, 2004.
- [2] Bederson, B.B., Quantum Treemaps and Bubblemaps for a Zoomable Image Browser (2001). *ACM Conference on User Interface and Software Technology (UIST 2001)* as PhotoMesa: A Zoomable Image Browser using Quantum Treemaps and Bubblemaps, pp. 71-80, HCIL-2001-10, CS-TR-4256, UMIACS-TR-2001-39.
- [3] Brown, Barry A. T., Sellen, Abigail J., O'Hara, Kenton P. (2000). A Diary study of Information Capture in Working Life. *CHI '2000*, 438-445. ACM Press, New York.
- [4] Chalfen, R. (1987). *Snapshot Versions of Life*, Bowling Green State University Popular Press, Ohio.
- [5] Ebay.com, <http://www.ebay.com>, Accessed on April 21, 2004.
- [6] Epinions.com, <http://www.epinions.com>, Accessed on April 21, 2004.



- [7] Flank, Sharon (2002). Multimedia Technology in Context. *IEEE Multimedia* 9(3): 12-17.
- [8] Fotonotes, <http://www.fotonotes.net>, Accessed on April 22, 2004.
- [9] Frohlich, David, Kuchinsky, Allan, Pering, Celine, Don, Abbe, and Ariss, Steven (2002). Requirements for Photoware. *Computer-Supported Cooperative Group '2002*, 166-175. ACM Press, New York.
- [10] The GEDCOM Standard Release 5.5, <http://www.gendex.com/gedcom55/55gctoc.htm>, Accessed on April 20, 2004.
- [11] Google Images FAQ, [http://www.google.com/help/faq\\_images.html](http://www.google.com/help/faq_images.html), Accessed on April 20, 2004.
- [12] Graham, Adrian, Garcia-Molina, Hector, Paepcke, Andreas, and Winograd, Terry (2002). Time as Essence for Photo Browsing Through Personal Digital Libraries. *Joint Conference on Digital Libraries '2002*, 326-335. ACM Press, New York.
- [13] Kules, Bill, Kang, Hyunmo, Plaisant, Catherine, Rose, Anne, & Sheiderman, Ben. Immediate Usability: Kiosk Design for a Community Photo Library. Department of Computer Science, Human-Computer Interaction Laboratory.
- [14] Ladd, Andrew M., Marceau, Guillaume, Bekris, Kostas E., Kaviraki, Lydia E., Rudys, Algis, & Wallach, Dan S. (2002). Robotics-Based Location Sensing using Wireless Ethernet. *MOBICOM '2002*, September 23-28, 2002, Atlanta, GA, 227-238. ACM Press, New York.
- [15] Lenat, D. (1998). *The dimensions of context-space*, Cycorp technical report, [www.cyc.com](http://www.cyc.com).
- [16] Lieberman, Henry, Rosenzweig, Elizabeth, & Singh, Push (2001). Aria: An Agent for Annotating and Retrieving Images. *IEEE Magazine*, July 2001, 1-6.
- [17] Liu, Hugo, and Lieberman, Henry. Robust Photo Retrieval Using World Semantics. MIT Media Laboratory, Software Agents Group.
- [18] Mills, Timothy J., Pye, David, Sinclair, David, & Wood, Kenneth R. Shoebox: A Digital Photo Management System. AT&T Laboratories Cambridge.
- [19] Ofoto.com, <http://www.ofoto.com>. Accessed on April 21, 2004.
- [20] Opensource Initiative, <http://www.opensource.org>, Accessed on April 20, 2004.
- [21] Rodden, Kerry, and Basalaj, Wojciech (2001). Does Organization by Similarity Assist Image Browsing? *SIGCHI 2001*, 190-197. ACM Press, New York.
- [22] Rodden, Kerry and Wood, Kenneth R. (2003). How Do People Manage Their Digital Photographs. *Proceedings of CHI 2003*, 409-416. ACM Press, New York.
- [23] Shen, Chia, Lesh, Neal B., Vernier, Frederic, Forlines, Clifton, and Frost, Jeana (2002). Sharing and Building Digital Group Histories. *Computer-Supported Cooperative Group '2002*, 324-333. ACM Press, New York.
- [24] Shneiderman, B. (2002). *Leonardo's Laptop: Human Needs and the New Computing Technologies*, MIT Press, Cambridge, MA and London, England.
- [25] Shneiderman, B., Kang, H. (2000) Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos. Proc. International Conference Information Visualization (IV2000), London, England.
- [26] Shneiderman, B., Kang, H., Kules, B., Plaisant, C., Rose, A., Ruchier, R. (2002). A Photo History of SIGCHI: Evolution of Design from Personal to Public. *Interactions* (9) 3, 17-23.
- [27] Singh, P. (2002). The public acquisition of commonsense knowledge. *Proceedings of AAAI Spring Symposium: Acquiring (and Using) Linguistic (and World) Knowledge for Information Access*. Palo Alto, CA, AAAI.
- [28] Snapfish.com, <http://www.snapfish.com>. Accessed on April 21, 2004.
- [29] Tan, Tele, Chen, Jiayi, Mulhem, Philippe, and Kankanhalli, Mohan (2002). SmartAlbum – A Multi-Modal Photo Annotation System. *Multimedia 2002*, 87-88. ACM Press, New York.
- [30] Toyama, Kentaro, Logan, Ron, Roseway, Asta, and Anandan, P. (2003). Geographic Location Tags on Digital Images. *Multimedia '2003*, 156-166. ACM Press, New York.
- [31] Wenyin, Liu, Dunais, Susan, Sun, Yanfeng, Zhang, HongJiang, Czerwinski, Mary, & Field, Brent. Semi-Automatic Image Annotation. Microsoft Research China.
- [32] Wenyin, Liu, Sun, Yanfeng, & Zhang, Hongjiang. MiAlbum – A System for Home Photo Management Using the Semi-Automatic Image Annotation Approach. Microsoft Research China.
- [33] Wikipedia.org, <http://www.wikipedia.org>, Accessed on April 21, 2004.
- [34] World-Wide Media eXchange, <http://wwmx.org>, Accessed on April 20, 2004.
- [35] World-Wide Media eXchange Download Page, <http://wwmx.org/Download.aspx>, Accessed on April 20, 2004.
- [36] Zhang, Lei, Chen, Longbin, Li, Mingjing, and Zhang, Hongjiang (2003). Automated Annotation of Human Faces in Family Albums. *Multimedia 2003*, 355-358. ACM Press, New York.
- [37] Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face Recognition: A Literature Survey in *ACM Computing Surveys* Vol. 35 No 4, December 2003, 399-45