"Three backups is a minimum": A first look at norms and practices in the digital photo collections of amateur photographers

Kristina M. Spurgin[*]
School of Information and Library Science
University of North Carolina at Chapel Hill

June 2, 2009

Serious amateur photographers are largely invisible in the literature on the management of personal photo collections. Leisure practitioners have been of little interest to researchers studying personal collections and the management thereof. The little that is known about leisure-related information behavior suggests that people develop and manage rich personal collections in the pursuit of serious leisure activities. Further, it appears that studying these collections has high potential for advancing research in the areas of personal digital collections, personal information management, and library and information science (LIS) in general.

Bodies of literature exist on the character, management, and use of photo collections by a) institutions and organizations; b) casual snapshooters collecting mainly family photos; and c) unknown individuals sharing, tagging, and annotating photo collections on the Web. Amateur photographers occupying a space sharing fuzzy boundaries with each of these areas. Gaining knowledge of amateur collections and collection-related practices will begin to fill in that space.

This chapter describes first steps exploring these potentials in the domain of amateur photography. The goals of the study are:

1. Answer the research question: Is there anything new and important to learn from studying the personal collection management of amateur photographers? If so, on what aspects of their practices should next efforts focus?

2. Answer the research question: What are the norms and expectations of the social world of photography regarding photography-related collection management?

[*]kspurgin@email.unc.edu, kristina@infomuse.net (permanent email).

3. Develop a qualitative data analysis template to be used in future interview-based research on this topic.

This chapter reports on preliminary analysis of a year of discussion threads from an online forum devoted to discussing collection management and digital photography workflow. Themes emerging from the data so far suggest that further study of amateur photographers' collection management practices will indeed uncover rich practices heretofore invisible in the gap between institutional/organizational, casual family, and publicly shared online photo collections.

The remainder of this chapter is organized as follows:

- Section 1 defines terms and explains the rationale for this work.

- Section 2 is a description of the methods used and my methodology, or reasons for using them.

- Section 3 reports on the topical and authorship characteristics of the forum studied.

- Section 4 explores broad themes of personal collection management in the forum discussions.

- Finally, section 5 discusses collection management practices described on the forum.

1 Background

1.1 Serious leisure

Serious leisure is a theoretical typology of approaches to leisure activities developed by Robert Stebbins in the context of his Serious Leisure Perspective. In the Perspective, serious leisure is differentiated from two other approaches to leisure: casual and project-based. Note that the Perspective is not a classification of leisure activities; rather it classifies different patterns of engagement with any particular leisure activity. For example, different people may approach photography as a casual activity, part of a project, or a serious pursuit. The Perspective's well-defined descriptions of leisure approaches facilitates clear communication about what kinds of leisure contexts are studied.

Serious leisure is of three types: volunteering, hobbies, and amateur pursuit. It is differentiated from casual or project-based leisure by its six distinguishing characteristics: (1) development of a *leisure career* in the chosen pursuit; (2) accumulation of knowledge and skills via *significant personal effort*; (3) *perseverance* at the activity, even when it is challenging, frustrating, or unpleasant; (4) a *social world and a unique ethos* forms around the activity; (5) *strong identification* with the serious leisure pursuits; and (6)

the enjoyment of eight durable benefits beyond the typical enjoyment found in leisure. The eight benefits may or may not be initial motivations for serious participation in an activity, but they are outcomes of participation. They are: (1) self-actualization; (2) self-enrichment; (3) self-expression; (4) regeneration or renewal of self; (5) feelings of accomplishment; (6) enhancement of self-image; (7) social interaction and belongingness; and (8) lasting physical products of the activity (Stebbins 2007).

An amateur pursuit is a serious approach to some activity, for leisure, that others pursue professionally (Stebbins 1998). Photography is typically pursued as an amateur activity. The link between amateur and professional means that the social world of amateur photography is part of a larger social world where expectations and standards of excellence are set by professionals. In the Serious Leisure Perspective, amateurism implies a striving for professional-level results, not unskillfulness. Amateurs are dedicated and critical members of the professionals' publics (Stebbins 1982). Some amateurs move into professional pursuit of their chosen activity by making it their job and main source of income. Other amateurs remain leisure participants; they may even make some money on their activity, but they do not depend upon that money as income.

1.2   Why study leisure contexts?

Relatively little is known about positively construed, non-work information behavior. Most research in LIS focuses on work contexts and problematic life situations (Kari and Hartel 2007). The number of studies inquiring about information behavior outside the workplace, and in unproblematic—even pleasurable—contexts increased in the past two decades, but it still makes up a small fraction of work in our field (McKechnie et al. 2002).

Research efforts should be directed toward understanding the informational aspects of the *higher things* of life, including serious leisure activities (Kari and Hartel 2007). Serious leisure is a form of lifelong learning (Jones and Symon 2001), an activity of great interest to libraries and their funding agencies (Urban 2007). Serious leisure participants become experts in their areas of interest. Much can be learned from self-motivated experts with successful practices, and from contexts where things are going right (Kotro 2007; Redmiles et al. 2005; Torrey et al. 2007).

A number of the existing studies on information behavior in leisure contexts have introduced findings that run counter to the "common knowledge" in our field. The common knowledge about personal collection management paints it as difficult, drab, and undesirable. Jones (2008) briefly describes "information warriors" who proactively take on PIM tasks without being triggered by negative events; however, most

3

PIM research finds that, for most people, information management is at best a low-priority tasks that is indefinitely postponed (Barreau 1995; Boardman and Sasse 2004; Jones 2008; Kaye et al. 2006). At worst, it is characterized as a tedious, unpleasant that should be eliminated (Cutrell, Dumais, and Teevan 2006). The same pattern is seen in home mode photography, where people report intentions of annotating photos and organizing them into albums, but never quite get around to it (Frohlich et al. 2002; Kirk et al. 2006; Miller and Edwards 2007).

Information management is viewed differently in some leisure contexts. Managing and making sense of found information is a core geneaology activity (Lambert 1996; Yakel 2004). Maintenance and systematic organization of collections is part of many collecting hobbies; the satisfaction of imposing order on some segment of the world is one attraction of serious collecting (Gelber 1999). Hobbyist gourmet cooks maintain personal culinary libraries requiring routine upkeep; their attitudes to this task range from "unconscious or nonchalant" to "zealous." In large personal culinary libraries, the cook "takes on the sensibility of a trained librarian" (Hartel 2007, p. 197–198).

Leisure participants also create information artifacts for other participants. This requires gathering and organizing information. In fandom, enthusiasts produce information objects organizing information about the phenomenon of interest: bibliographies, directories, chronologies, discographies, catalogs, and listings (Hart, Schoolbred, and Butcher 1998). Computer and electronics hobbyists spend significant amounts of time and effort producing how-to documentation to share online (Torrey et al. 2007). Elementary school children create and share hobby or leisure-related information for fun (Trace 2008, p. 1550). Information creation and sharing remains a fundamental aspect of leisure activities such as fan fiction (Kustritz 2003; Thomas 2007) and role-playing games (Hammer 2007).

1.3   Amateur photographers

In this chapter, the term *amateur* refers to the approach to a leisure activity as described by the Serious Leisure Perspective. This distinction is necessary because existing literature imprecisely uses the phrase *amateur photographer* to refer to home mode photographers and camera club members.

The home mode is a pattern of interpersonal and small group communication centered around the home and focused on pictorial materials (Chalfen 1987). Home mode photography has no professional counterpart. It documents other leisure pursuits rather than being pursued itself as a leisure activity. In the Serious Leisure Perspective, it would be considered a casual pursuit. In the literature of art, sociology, and photowork/PIM, the home mode is variously referred to as *domestic*, *amateur*, *family*, *vernacular*, or

snapshooter photography.

Camera clubs have existed since soon after the discovery of photography, and have developed their own distinct practices and aesthetic (Grinter 2005; Schwartz and Griffin 1987). Though there is no direct linkage between camera club photography and any particular flavor of professional photography, it is included as a subcategory of amateur photography here for two reasons. First, a dynamic exists between the clubs and the camera and film industry as commodity agent, and has been present since the early days of camera clubs (Griffin 1987; Slater 1991). A commodity agent is a group or individual "involved in the production, facilitation, and exchange of activity related commodities" (Yoder 1997). Stebbins now includes these actors in the complex dynamic between amateurs and professionals (Stebbins 2007). Second, aspects of the camera club aesthetic have many traits of landscape, wildlife, portrait, and some advertising photographies, though there is not a direct connection between those professional practices and the practices of camera clubs.

The existence of separate social worlds of professional art, advertising, and photojournalist photography Rosenblum (1978) suggests the existence of amateur photographers outside the camera club, yet free range amateurs have until recently been invisible in the literature. Four recent studies of the photo-sharing website Flickr.com describe some participants that seem to fit the profile of free range serious amateur without clearly defining them as such. These studies and the amateur traits mentioned in each are summarized below. For the purposes of this study, I assume that the participants described below are indeed serious amateurs.

(Davies 2006) – Based on email questionnaires sent to the author's Flickr contacts and her own experiences and observations on Flickr, this article examines the site as an environment for learning and teaching. Some users engage in in-depth social learning about photography, with long-term increase in skills and confidence. This is an example of the acquisition of knowledge and skills through effort, with the durable benefits of self-image enhancement and feelings of accomplishment. Some users develop new ways of seeing through participation in the site. Learning to see like a photographer is a key part of taking on the role of photographer (Rosenblum 1978). Finally, Davies reports on the emergence of shared social procedures and values—an indicator of the presence of a shared social world, which can also lead to social interaction (online or off) and a sense of belongingness.

(Burgess 2007) – Burgess studied Flickr group members' use of new media to articulate vernacular creativity. Some participants were interested in professional standards for photography and devoted significant time and effort to learning about advanced camera techniques and the theory of photography. One participant sought formal photography training. This recalls the amateur's dedication to acquisi-

5

tion to knowledge and skill as well as his link to the photography professions.[1] Participants had arcs of increasing interest, knowledge, and skill in photography (leisure careers). These were marked by the purchase of digital single lens reflex cameras (DSLRs), investment in lenses, and a desire to regularly upgrade their equipment. This pattern of consumption is characteristic of serious leisure (Stebbins 2007). Participants described their photography in terms of art and self-expression (a durable benefit), and reported forming social groups based on photo activities. Burgess identified social worlds of photography with aesthetics, discourses, and best practices. Professionals non-professionals mingled in these social worlds, and professional values and standards ruled.

(Miller and Edwards 2007) – This study unexpectedly turned up a group of participants (all members of one Flickr group) with a different approach to photography and photo sharing. The paper calls them *Snaprs*. These users took photos more frequently, and processed them more promptly than home mode photographers. Snaprs considered photography a hobby, were more technologically literate, and expressed no concern about the privacy of their photos. They shared their photos publicly and tagged photos to increase their visibility.

(Cox, Clough, and Marlow 2008) – This paper refers to serious leisure in its title and claims to examine Flickr use in the context of the Perspective. However, the researchers demonstrate a questionable understanding of the Serious Leisure Perspective as a framework by describing *serious amateur photographers* in the context of serious leisure and then referring to their photo activities as a *hobby*. The paper identifies the activities of traditional *hobby* photography as including purchasing equipment, developing social relationships around photography, and an emphasis on learning. Another activity mentioned is exhibiting photos and participating in competitions. Exhibition is one of the notches in the belt of an art photographer that can be used as a signifier of professional status (Rosenblum 1978).

<div align="center">⊱ ⊱ ⊱</div>

Services like Flickr may simply be increasing the visibility of the existing free-range, solo amateur photographer. An alternate view is that such services are fundamentally changing the landscape of amateur photography by introducing new affordances (Cox, Clough, and Marlow 2008), eroding or redrawing existing boundaries between kinds of amateur photography, and blurring the line between amateur and professional (House 2007; Murray 2008).

In summary, amateur photographers, as defined in the Serious Leisure Perspective, include the proto-

---

[1] Amateur photography is a leisure activity pursued by men much more heavily than by women (Cox, Clough, and Marlow 2008; Slater 1991). Based on this, I've default to masculine pronouns unless I'm referring to someone clearly identified as female on the forum.

typical camera club member, as well as an array of other serious non-professional photographers. Amateur photography in this chapter does not include the casual, consumer, family photographers often called amateurs in the literature. Those photographers are referred to here as home mode photographers. For the purposes of this chapter, the possible amateurs described in the four Flickr studies are assumed to be amateur photographers. Below, some findings from these studies will be compared with themes and practices described on the forum.

1.4   Why study amateur photographers?

The choice to study amateur photographers was motivated by the fact that knowledge about their collection management practices will fill a gap in the literature on image organization and retrieval. This group is also interesting because managing photographs presents unique challenges.

Much work has been done on the collection, organization, description, and retrieval of images in cultural heritage, organizational, and educational settings (Anderson et al. 2006; Baca 2002; Chen and Rasmussen 1999; Jorgensen 2003; Neal 2006; Teper 2004). A respectable body of work also exists on how home mode photographers use and manage photographs (Adams, Cunningham, and Masoodian 2007; Frohlich and Fennell 2007; Kirk et al. 2006; Nightingale 2007; Rodden and Wood 2003). Finally, the literature on the use of online tools (primarily Flickr) for sharing and tagging photographs is rapidly expanding. (Angus, Thelwall, and Stuart 2008; Cox, Clough, and Marlow 2008; Rafferty and Hidderley 2007). Results of research on amateur photographers can be linked to findings in these three areas because amateur photographers are a) individuals working with personal collections; b) working with often vast collections, at least part of which are made public; and c) using the web to share photos.

The complexity of providing intellectual access to images is well-documented (Jorgensen 1999; Krause 1988; Layne 1994). The attributes that may be salient for retrieval are numerous and layered. Multiple approaches and facets of image subjects exist. Unlike text documents, photos cannot yet be automatically organized and retrieved semantically. Very recently, some personal photo management applications have added face recognition features (Apple 2009; Google 2009); however, this sort of content based image retrieval is a relatively young technology and we are far from bridging the *semantic gap* using automated techniques (Enser et al. 2007).

7

1.5 Summary

Studying personal collections in leisure contexts is needed in order for LIS to gain an holistic understanding of information practices. Knowledge about how self-motivated, self-supporting experts understand and work with information in their pursuits may help libraries meet the needs of lifelong learners and those with more casual interests in leisure pursuits. Amateur photographers are an important group to study because the management of digital photos presents unique challenges. In an increasingly visual culture, information professionals much understand photograph collections and how people work with them (Marcum 2002). Amateur photographers have not been studied in LIS, and so there is a knowledge gap to be filled.

2 The Study

This chapter reports on an exploratory qualitative analysis of a year's worth of threads on a photo.net forum devoted to discussing digital photograph collection management and digital photography workflow. The purposes of the study are: 1) to assess the likely contribution of further research into amateur photographers' collections; 2) to learn about norms and expectations concerning digital collection management in the social world of photography; and 3) to develop a qualitative data analysis template to be used in future interview-based work on the topic. This section describes the methods I used and my reasons for using them. At the end of the section, I discuss limitations of the design.

2.1 Methodology

An exploratory qualitative analysis of a well-known photography forum was conducted because it allowed initial forward movement on each of the research goals to proceed simultaneously. As the forum threads already existed, it was possible to survey them for the presence of topics and practices in the community's discourse without influencing responses. The presence of previously unseen personal collection management practices and concerns was assumed to be indicative of phenomena of research interest and promising for future research.

Amateur photography is part of a social world (Unruh 1980) in which standards of excellence are set by professional photographers. The dispersed nature of this world requires reliance on mediated communication in order to transmit, maintain, and develop the discourse, values, and norms of the social world. Online photo sharing services like Flickr provide mediated spaces for social learning about *seeing like*

a photographer and technical photography skills (Burgess 2007; Davies 2006). It follows that a primarily text-based service focused on collection management issues provide space for social learning about the norms surrounding collections and their management. The data represent a range of photographers larger and more diverse than could be personally surveyed, so topics and practices could be seen across numerous individuals.

Future interview research will be conducted with amateur photographers, and a template analysis of that data is planned. Template analysis is a good approach for an unexplored topic with a large body of closely related literature because it lies between the extremes of grounded theory and content analysis (King 1998). Use of the template in the current study will show whether topics and themes from the literature represented in the template map to real-world issues in the personal collections of amateur photographers. The template can be augmented, refined, and shaped to more realistically reflect the collection management issues faced by amateur photographers.

2.1.1 Limitations

The methods used in this work have serious limitations. The use of one online community potentially biases the results. Voices of those who do not follow the norms of the forum, or who do not have problems or strong opinions may be silent. The use of forum data means that no follow-up questions could be asked to clarify the meaning of ambiguous or confusing statements. There was no way of knowing who all of the forum authors are, or what their motivations for authoring threads and comments are. Each thread is a vignette torn from an unknown, richer context.

Using an exploratory analysis to develop a template requires no tests to demonstrate the validity or reliability of coding—that step is carried out during the analysis that uses the template developed here; therefore, the trustworthiness of the present analysis cannot be formally demonstrated. To mitigate this, an audit trail is provided for statements based on the data. The unique ID of the ATLAS.ti quotation(s) related to the statement is inserted in square brackets. The data are available from the author in XML or native ATLAS.ti .hu format.

In no way are the themes reported here generalizable. They exist in the social sub-world of the forum examined. These findings encourage future research into whether these themes run through the broader world of amateur photography.

2.2 Methods

2.2.1 Source selection

Photo.net was selected as the data source based on the many references to the site encountered in informal exploration of the online world of amateur photography and personal awareness of several amateur photographers active on the site. The Photo.net forums capture candid communication between professional and amateur practitioners sharing information pertinent to real-life photography situations and concerns.

In addition to its 33 active forums devoted to a range of photography topics, Photo.net includes articles, galleries, and reviews. An *About Photo.net* blurb on the site's home page contains several indicators of serious leisure/professional orientation (emphasis added):

> photo.net is a site for **serious photographers** to **connect with other photographers**, explore photo galleries, **discuss photography**, **share and critique photos**, and **learn about photography**. The site began in 1993 as Philip Greenspun's personal home page at MIT and has grown to become **a community of photographers** that includes more than 675,000 registered users **working to help each other improve**. Our editorial goal is to serve busy readers who want clear answers to questions. When we review equipment, we do it from the perspective of "In what kinds of photographic situations would this be **the right choice**?" In **building community**, we value members who provide **constructive criticism** and **helpful assistance** to other members.

First, the audience is identified as *serious photographers*. Social connection and desire to learn and improve skills are emphasized. Discussion and critique of photography build and transfer a shared discourse across the community. The mention of *right choices* implies that there are wrong choices, suggesting some level of shared community values, standards, and practice. Finally, while community building and social networking are a goal of the site, the busy users are not coming to the site to hang out and collect friends—the focus remains on efficient, useful sharing of information.

2.2.2 The forums

Most of the activity on Photo.net takes place on 35 topical forums. The home page claims that over 3,000 new threads are added daily. The forums are placed into the five broad forum categories listed below:

- **General photography**. 6 forums.

- **Image related**. 3 forums.

- **Photography practice & technique**. 10 forums.

- **Photography equipment**. 15 forums.

- **Other**. 1 forum.

2.2.3   The sample

The Digital Darkroom forum is the focus of this study because much of its content is directly related to personal digital collections. This forum is in the Photography Practice & Technique category. The archived threads in the Digital Darkroom forum are divided into 40 thread categories. Fourteen (14) thread categories were selected as the most directly relevant to personal digital collection management. The thread categories are discussed in more detail in section 3 and are summarized in Table 1. The sample was defined as a year's worth of threads in order to avoid bias by seasonal fluctuation of activity or themes. The sample was defined as all threads posted in these 14 thread categories between 29 January 2008 and 29 January 2009 (n=1008).

2.2.4   Data acquisition and processing

Scraping data from the site was easy due to the structure of the forums. The free software package Wget was used to download threads from the relevant categories (Nikšić 2007). Wget was set to request one post every 45 seconds in order to comply with Photo.net's Terms of Use (Root 2008). The site scrape resulted in the download of 1008 separate HTML files over three days.

These files were concatenated into 26 large text files, each containing the HTML of up to 50 threads; HTML files from different thread categories were lumped into separate text files. A series of Emacs macros was used to clean the 26 text files. The HTML and JavaScript code from the photo.net web pages was stripped; it was replaced with simple markup tags to retain the structure of the data.

2.2.5   Structural coding and analysis

Each of the 26 text files was copy/pasted into an ATLAS.ti free memo as plain text. The memos were assigned as primary documents (PDs). PDs from the same Photo.net thread categories were grouped into PD families to allow for limiting queries to specified thread categories. ATLAS.ti's autocoding feature was then used to quickly segment the data based on its structure. Each thread and its followup comments became an ATLAS.ti quotation coded as a THREAD. Likewise, each original thread posting (minus comments) was coded as an ORIG-THREAD.

For a broad view of activity on the forum, basic patterns of thread and comment authorship were analyzed. Photo.net requires users to register with a unique user name in order to add forum content; therefore, an assumption is made that a user name represents one author, and that one author uses one user name. The names of thread and comment authors were extracted from the structured text and pasted into an Excel worksheet, where pivot tables were used to summarize the number of threads and comments authored by each registered user appearing in the sample. Results are discussed briefly in Section 3.2.

2.2.6 Template development through qualitative analysis

The initial template was constructed before data collection began. It was based on the literature of personal collections, photowork, serious leisure, PIM, and image description and was divided into five key areas, as described below.

- **Contextual aspects,** including photographic activities, types of information in the collection, social aspects, the role of physical space (and conceptions of digital space), and tools and services used.

- **Organization and maintenance activities,** including collection organization strategies; metadata creation, manipulation, and use; collection maintenance strategies; and collection storage strategies.

- **Collection maintenance evaluation criteria**, including definitions of concepts such as *archiving* and *backup*, and the characteristics on which collection management-related tools, strategies, and practices are judged.

- **Collection maintenance situations**, or the types of situations in which amateur photographers engage in collection maintenance such as organizing, weeding, backing up, or migrating. This area also includes the kinds of steps taken to develop strategies when in these situations.

- **Influences**, or the objects, events, people, named as helpful or unhelpful in collection maintenance.

This template was then applied to the forum data. Where topics and practices fit neatly into the template, they were coded. Where they did not, but there was a clear place in the template for them, the template was expanded to accommodate them. If the template did not have a clear place for them, a note was made. A frequently occurring uncodable topic or issue was taken as evidence that the template did not fit the reality of the community and the template was re-organized. This was usually minor, but two major changes were made to the template. These are mentioned where appropriate below.

12

To support focused and systematic analysis, I asked the following questions in roughly the same order for each original thread.

- What is the situation, question, or problem that brought the author to post this thread? How do they characterize it?

- What steps have they taken toward moving through that situation by posting this thread? For example, did they ask for explicit factual information, procedural information, technical support, opinions, or reports on the experiences of others?

- What collection maintenance activities or strategies are named or implied?

- What kinds or formats of information are named?

- What tools or services are mentioned in relation to managing this information?

- What evaluative statements are made regarding these tools, services, information types, strategies, and/or activities?

- What is mentioned as being helpful? A hindrance?

The preliminary qualitative analysis focused on the original threads posted to the forum. Comments on some threads of particularly direct relevance were analyzed. For example, responses to a question about how people organize their photo directories were coded. Just under 50% of all original threads were included in the analysis presented here.

3 The Digital Darkroom forum

3.1 Thread categories

Table 1 displays the Digital Darkroom forum thread categories selected for inclusion in this study. The table should be read as follows: 1169 threads have been posted in the Computers category over all time (# in for). This category contains 24.9% of all of the threads ever posted in the Digital Darkroom form (% of for). 300 threads from the Computers category were included in the sample (# in samp). That makes up 29.8% the sample (% of samp). 25.7% of all threads ever posted to this category are included in the sample (% of cat). Finally, 66.6% of the sampled threads in the Computers category were included in the

analysis reported in this chapter (% rep).[2] The table is organized in descending order by percentage of sample falling into each category.

photo.net label	study label	# in for	% of for	# in samp	% of samp	% of cat	% rep
Computers	Computers	1169	24.9	300	29.8	25.7	66.6
Software—Utilities	Utilities	600	12.8	128	12.7	21.3	0
Software—File Formats	Formats	582	12.4	123	12.2	21.1	33.3
Archiving	Archiving	491	10.4	108	10.7	22.0	66.6
Storage	Storage	446	9.5	82	8.1	18.4	0
Images—Libraries	ImageLibs	210	4.5	59	5.9	28.1	100
Software—Capturing	Capture	280	6.0	49	4.9	17.5	0
Software—Web Publishing	WebPub	191	4.1	36	3.6	18.8	0
Books	Books	149	3.2	34	3.4	22.8	100
Services—Printing	Printing	189	4.0	27	2.7	14.3	100
Software—Archiving/Billing	ArchBillApps	128	2.7	23	2.3	18.0	100
Printing—Archival	ArchPrinting	168	3.6	20	2.0	11.9	100
Migration	Migration	70	1.5	16	1.6	22.9	100
Services—Hosting	Hosting	29	0.6	3	0.3	10.3	100
	Total	**4702**	**100.0**	**1008**	**100.0**		**49.1**

Table 1: Summary of thread categories included in study

Forum authors interpret the boundaries of these categories loosely; similar threads are found in different categories. Though the categories for these topics have not been analyzed, many analyzed threads are about choosing the right software, planning storage and backup systems, and what web services are the best. The forum and the sample are dominated by the Computers category. Since post-shoot digital photography activites are largely conducted on computers, the topic is even more dominant than the numbers indicate.

3.2 Thread authorship

Just under half of the thread content has been analyzed, but the authorship analysis included all threads from the year. During the year, 718 unique authors (user names) created 1008 threads. Authors created between 1 and 23 threads each. Only 5 authors created more than 6 threads. 572 authors created only one thread. The frequency distribution of number of threads created per author exhibited a Lotka's law-like shape: as the number of threads created by an author increases, the number of authors creating that number of threads decreases.

The same basic pattern emerged among commenters. 1404 unique commenters left 6872 comments. Authors created between 1 and 324 comments each. Only 30 authors created 30 or more comments. Only

---

[2]Digital Darkroom thread categories excluded from the study were: Color Management—Calibration, —Color Space, —Lighting; Imaging Techniques—Emulating (B&W, IR, filters, etc.), —Other, —Resizing/Sharpening, —Retouching; Monitors; Printing—B&W, —Inkjet Inks/Papers, —Other, —Technique, —Printers—Home, —Printers—Pro/High volume; Projection; Scanning—Other, —Scanners, —Software, —Technique, —Scanners—Drum, —Scanners—Film, —Scanners—Flatbed; Services—Educational, —Scanning; Software—Editing, —Stitching, HDR, compositing.

18.95% of authors (n=266) created 5 or more comments. Nearly half (49.29%) of authors (n=692) made a single comment. Again, the distribution took the shape of a power function where as the number of comments made by an author increased, the number of authors making that many comments decreased.

The same person posted the largest number of threads and comments. Four other authors were in the top 30 most prolific authors and commenters. This suggests there is a very small group of core contributors to the forum. The majority of its content is written by a very large number of infrequent or one-time contributors, at least some of whom are regular lurkers [22:56, 12:44]. The forum is a mediated communication tool for sharing knowledge the dispersed social worlds of photography.

While most authors do not explain their interest or involvement in photography, some do identify themselves as professionals or amateurs. Presence of both on the forum confirms that this is part of a photography social world with the Professional–Amateur–Public dynamic characteristic of amateur pursuits. The results of this study are not only about amateur photographers, then, but about serious photographers. Surely some non-serious photographers are present, but the working assumption is made that active authors are predominantly serious photographers. Results cannot be extended to all serious photographers but are evidence of themes present in one community of serious photographers that might fruitfully be looked for in another.

4   General themes

What emerges from the analysis thus far is a blend of confirmations of existing knowledge about how people work with photo collections, contradictions to that knowledge, and new themes. This section discusses forum themes related to some of the over-arcing issues commonly discussed in the literature, as well as some new themes.

4.1   Functions of personal collections

The functions of personal collections include: (1) to support sharing of collection contents; (2) to assuage fears of loss; (3) to support re-finding; (4) identity construction; and (5) bearing witness to a legacy (Kaye et al. 2006; McKemmish 1996; Van Dijck 2005; Williams et al. 2008). The Serious Leisure Perspective suggests a role for the latter two functions in the amateur collection, but forum authors do not talk in terms of identity construction or legacies.

Sharing is a central function of home mode photo collections (Adams, Cunningham, and Masoodian 2007; Frohlich et al. 2002; Kirk et al. 2006). The amateur photographers in the Flickr studies previously

15

discussed shared photos, but they did so more openly than home mode photographers. Four kinds of sharing are mentioned by forum authors: with known people [15:180, 19:104]; with the public via photo sharing sites or the web at large [27:76, 11:40]; with clients [16:87, 9:24]; and with the art world via exhibition [24:15, 30:30]. The presence of the latter two indicate the P–A–P dynamic of the photography social world.

People dread losing their home mode photos: "family photos" are top of the typical list of what to save from a home in a disaster. Ironically, their digital collection-keeping often contradicts their stated values on this matter. Many people back up their data infrequently or not at all(Kirk et al. 2006; Marshall 2008). The spectre of "losing everything" also haunts the forum [12:31, 34:73], with the difference that authors report going to extensive measures to guard against the possibility (see discussion of backing up in section 5.3).

In contrast with home mode photographers who exhibit little interest in retrieving or exploring old photos from their collections (Kirk et al. 2006), forum authors express the need for very specific search and filtering functions such as finding images by camera model [22:62, 22:224], focal length [11:143, 22:250], image resolution [22:114], and image ratings [22:16]. This is indicative of the forum's professional values: pros must be able to provide former clients with specific photos on demand. Also, the ability to find images created in a certain way, or of a particular quality, enables the use of the collection as a reflective tool for self-assessment and further learning [22:293].

4.2   Lack of clear guidelines for safekeeping

There are few sources of clear advice for forum authors regarding management of their digital photo collections for the long term. The standard source on this topic for forum authors is *The Digital Asset Management (DAM) Book* (Krogh 2006) (see section 4.8), but some of its advice contradicts existing preservation guidelines. Anderson et al. (2006) provides guidelines for institutional digital image collections and highlights gaps in our knowledge of how to best handle these materials. Paradigm (2008) contains an appendix of guidelines for creators of personal archives, but is not focused on the particular issues for such archives that consist predominantly of photographs. There is a lack of information about managing digital photo collections that is: a) targeted to serious photographers; b) shows understanding of and respect for the ethos of serious photography; and c) contains clear guidelines for digital photo preservation from authoritative preservation professionals.

4.3 Size and complexity

The bloated size and complexity of home mode collections can be attributed to the low cost of producing digital images and the ease with which they can be edited and duplicated. Once you own a digital camera, taking photos costs you nothing but digital storage space, which always getting cheaper (Beagrie 2005). Why not take "over 100 images of the same sheep on a hillside" [22:294]? What to do with the original file, the one with the touched up red eyes, the cropped one I sent to my parents, the low resolution one I uploaded to the Web? As Marshall (2008) points out, while lots of copies may keep things safe, they also make it even harder to keep track of what you have and what you don't.

The situation is more pronounced for the serious or professional photographer. Not only do they shoot more images more frequently, the images they create are much larger that those of the home mode photographer. This is due to their tendency to purchase ever-nicer cameras and other equipment [(Burgess 2007; Cox, Clough, and Marlow 2008; Stebbins 2007), as well as their general insistence on the highest possible quality. One author writes, "so my final image is saved as a TIFF no compression, in 16-bits mode, 360 resolution. i just finished working on an image in those settings, and the file size has come up to 59MB.. is this normal?" The first commenter's response: "Yes, that's normal. Hopefully you're also keeping your RAW[3] files" [22:144].

The availability of ever-larger storage devices at reasonable prices allows photographers to continue building their collections, even as file sizes creep upward; however, large files pose two further challenges. First, the file size of an image is drastically increased if certain color modes or processing/ editing methods (such as addition of layers in Photoshop). Second, powerful software is needed to work with very large image files. These applications are themselves huge, but more importantly, they require powerful hardware to run efficiently[4] This leads serious photographers to buy or build powerful custom computers dedicated to working with images. The extreme example was an author asking what kind of computer to put together to handle working with large format sheet film scans at 3000 dpi,[5] resulting in images file sizes in the range of 3 GB each [15:92].

Most of the Computers thread category is given over to this sort of question. The initial analysis template bracketed out this sort of technical discussion (and discussions of shooting and editing photos) because a) these topics are not discussed in the photowork literature; and b) they seemed too broad

---

[3] A RAW file is the original image data directly from the camera with no algorithmic corrections or compression applied.

[4] "Photoshop requires at least 1 GB of free hard-disk space, but more is recommended. If you have more than one hard disk volume, you should specify additional scratch disks. Photoshop CS2 supports up to 64 exabytes (EB) of scratch disk space on a total of four volumes. (An EB is equal to 1 billion gigabytes.) RAID 0 partitions provide the best possible performance as Photoshop scratch disks" (Adobe 2008).

[5] dpi = dots per inch, or image resolution. The higher dpi, the better print image quality.

and not directly related to collection management issues. But the importance of these relationships was strongly evident on the forum, particularly in the Computers thread category and in the many threads about *workflow* (discussed in the next section). Figure 1 illustrates the complexity of the relationships between these considerations.[6] Many of these complexities become apparent in the design and choice of workflows. Finally, the Paradigm guidelines for collection creators touch on each of these intertwined topics (Paradigm 2008). Based on this, I made a large change in the template by adding codes for describing hardware, software, and file formats.
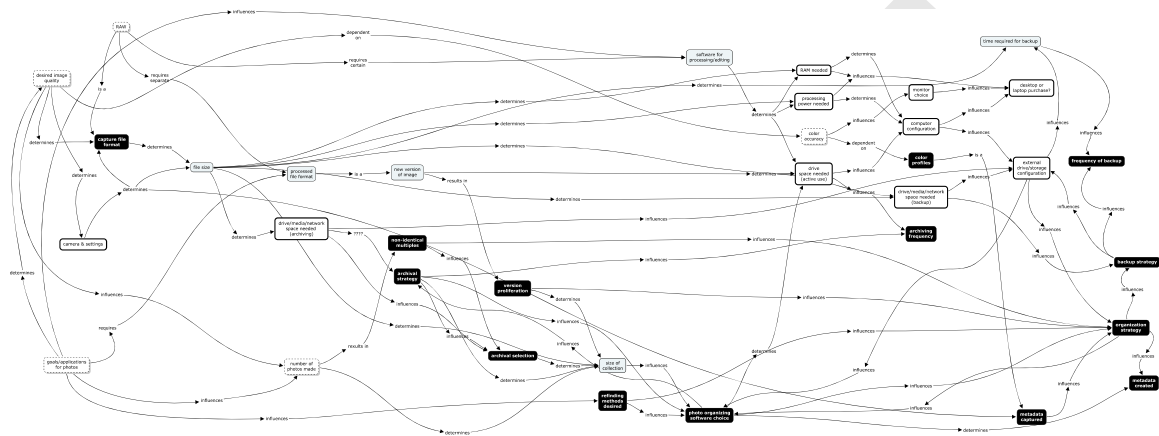

Figure 1: Tightly intertwined hardware, software, file format, and information management considerations. Black nodes are core collection management concerns. White nodes with thick black outlines are hardware issues. White nodes with dotted lines and a drop shadow are photography practices. The light grey nodes are properties and characteristics that do not fit into these categories.

4.4 Importance of workflow

A common forum topic is the development of digital photography workflows—all of the steps done to process and manage photos, from downloading from the camera, through backing up or archiving. Forum authors are concerned with having the "proper" or best workflow [19:63], where efficiency is a key criterion for evaluation. People ask for help and guidance in developing their own workflows:

> "Hopefully someone can recommend a good book on a proper and efficient workflow" [9:110].

> "I am planning on using Lightroom and wanted a book to walk me through an efficient workflow and for some tips regarding image processing" [9:129].

> "Is there some better workflow I'm not thinking of? What works for people?" [13:72]

---

[6]A larger, readable version of this graphic is available at http://infomuse.net/papers/idigital/figureA.pdf

One of the milestones of becoming a serious photographer is to learn to shoot and process RAW images; working with RAW files requires steps unnecessary when shooting JPGs (the default). The typical amateur or professional workflow is complex, uses a number of applications, and results in multiple versions of photos in various file formats. Figure 2 illustrates a relatively simple workflow for processing RAW photos; the author described his existing system, asked if it is optimal, and proposed a backup/archive strategy for comment [19:66].

Workflow is at the core of the amateur's personal digital collection management, where a structuration-like duality[7] is found between workflow and collection: workflow shapes and is shaped by the use and growth of the collection, while tools and services are chosen (or built) to fit into a workflow shape it via their affordances, functions, and limitations.
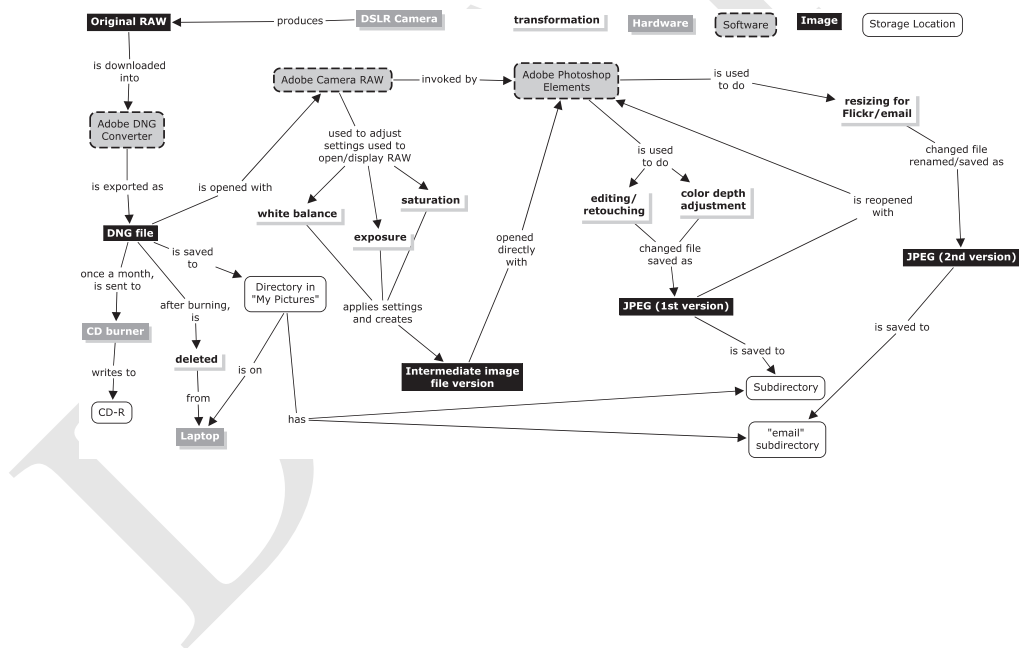

Figure 2: Example of relatively basic workflow (partially in practice, partially proposed by author) [19:66].

---

[7]See (Giddens 1984).

19

4.5 Proactive information management

As described above, many forum authors proactively assess, reassess, plan, design, redesign and tweak their workflows. Designing a workflow is a high level collection management task at the *meta-level* of personal information management (Jones 2008). Developing a good workflow requires consideration of several activities included in the Paradigm guidelines discussed below: naming and structuring files and folders, keywording/tagging or adding other metadata, choosing file formats and software, devising a backup strategy, and, to some extent, system administration and hardware/media management.

A number of authors posted on the forum because they decided to catalog their photo collections: "I am about to embark on the arduous task of organizing, cataloguing, and backing up several terabytes of digital images, and would like to get it right the first time" [11:77]. Authors ask for book and software recommendations as well as procedural information regarding the best ways to get organized [11:38, 12:51].

Finally, many authors write about how to build the best backup systems to avoid data loss. Making sure things are backed up before problems occur is the expected and approved behavior in this social world. Two authors volunteered that they had not backed up their data. One asked how to make backups using a particular tool [11:196], while the other was in a software malfunction/possible data loss situation, vowing to make backups in the future [22:19]. All except the latter are taking proactive steps to manage and maintain their collections before a problem strikes.

In contrast with the participants in many personal collection or PIM studies who do reactive collection management in response to problems, forum authors commonly engage in proactive collection management in response to breakdowns. This proactive approach requires intentional engagement with breakdown situations. The term *breakdown* is used here in a Heideggerian sense[8] as briefly summarized by Winograd and Flores (1986):

> A breakdown is not a negative situation to be avoided, but a situation of non-obviousness, in which the recognition that something is missing leads to unconcerning... some aspect of the network of tools that we are engaged in using (p. 165).

Forum authors engage breakdown situations in three main ways. First, some create breakdowns to get a clearer look at a specific aspect of a system. This manifests in setting up quality tests and doing benchmarking [9:54, 14:39]. Second, they anticipate possible breakdowns and plan for prevention and recovery. This is clearly seen in the attention given to designing backup systems. Finally, an amateur

---

[8] As elaborated in (Heidegger 1978).

career involves acquisition of new knowledge and skills through personal effort and perseverance; in other words, an amateur decides to walk a path he knows will be strewn with breakdowns, for the process of continual learning requires incessant pushing into spaces where the concepts and skills he needs are not ready-to-hand. For the amateur, part of the enjoyment of the activity is breaking down, figuring out how to get running again, and moving on to the next challenge.

The initial template conceptualized collection management activities taking place in positive, neutral, or negative situations. This coding structure soon proved too simplistic. The template must be revised to attend to subtler situation factors like (1) the authors' expectations and how their experience maps to expectations; (2) the authors' level of preparedness for the situation (including possession of resources: emotional, knowledge (including procedural or "tech skills"), time, access, and money); and (3) the authors' sense of control in the situation.

The current conclusion is simply that forum authors frequently mention doing proactive collection management activities. Collection management is part of the amateur leisure pursuit. This echos the Snaprs' enjoyment of organizing their collections (Miller and Edwards 2007), but stands in contrast to the findings of most photowork and PIM studies.

4.6   Technical knowledge and skill

One of the resources that increase preparedness for breakdown situations in collection management is adequate technical knowledge and skill. Previous studies have demonstrated that many who manage personal collections—photo or otherwise—display a lack of the most basic understanding of how computers work, how to use them, and the nature of their digital information (Marshall, Bly, and Brun-Cottan 2006; Williams et al. 2008). Often people rely on others to provide ad hoc technical support (Marshall 2008).

Many forum authors qualify their thread posts with admissions of not knowing much about computers or technical stuff [14:171, 14:180, 15:124, 15:242], but most of these are trying to learn more. So far, there is only one example of lack of basic computing concepts: installation of Photoshop on an external hard drive [22:101]. All of these appear to be operating at a higher level of technical knowledge than described in (Williams et al. 2008). The forum itself is a means of getting technical support from others, and some forum authors mention that they receive offline ad hoc tech support from others [12:77, 15:119, 19:129].

It is common, however, for forum authors to mention that they provide technical support for others. One author does tech support at a summer camp [22:200]. Others mention that they are posting to

get answers on an issue they are working on for someone else: a girlfriend [27:38], a friend [13:67], and parents [12:118, 11:17, 15:235]. One person mentions that he organizes other people's photos for them, but does not explain in what capacity he provides this service [22:290]. Overall, as in (Miller and Edwards 2007), the technical proficiency of forum authors seems be higher than average.

The high level of technical competence shown on the forum is in line with the characteristics of amateur photographers. From the time of the discovery of photography, they have been interested in the technical details of photographic processes. At first this required knowledge of how to prepare plates and work finicky camera equipment. Extensive knowledge of chemistry and properties of light remained key in print photography. Now the necessary high technical skills of the professional and amateur are computer skills.

4.7 Conceptual ambiguity

Williams et al. (2008) found "a certain ambiguity" in use of the terms *back-up*, *storage*, and *archive* by the creators of personal collections. This ambiguity is also noted in (Marshall 2008) and is present on the forum. A few authors clearly use *backup* to refer to regular duplication of current files to enable recovery of active documents [11:58, 16:146, 22:240]. Others suggest archiving is for long-term storage [12:114, 12:158]. Most authors do not clearly indicate what they mean by these terms, which are used rather loosely and interchangeably, such as in the request for recommendations of "archiving software for incremental backup of image files" [11:145] and the question of whether labeling DVDs would affect the "archivability" of the backups [12:66]. The forum is an inappropriate source of data on which to base understanding of serious photographers' conceptual understandings of these terms, so the remainder of this chapter makes no distinctions among them. Practices related to these terms are summarized in section 5.3 on page 33.

4.8 A unique ethos

A social world and a unique ethos will form around a serious leisure activity. In the case of an amateur activity, this social world and ethos will be embedded in a Professional-Amateur-Public dynamic informed by professional standards and values (Stebbins 2007). Threads on the Digital Darkroom forum suggest it functions as a mediated communication space for members of the dispersed social world of photography. Forum authors include self-described professionals, semi-pros, and amateur photographers, as well as many who do not describe their level of photography activity. Forum content suggests the existence of a

shared set of values in three ways: (1) one book appears to be the authoritative text on managing digital photo collections; (2) norms and best practices are implicitly and explicitly referenced; and (3) common practices and advice emerge.

The forum includes no formal guidelines for managing collections, but *The DAM Book* (Krogh 2006) is considered the most authoritative print source on the matter. In some threads, authors mention that they have already read this book, as though doing so were expected or commonly recommended. Others base their systems on the book's suggestions, or plan to do so [37:72, 12:168]. In the comments that have been analyzed, commenters frequently advise thread authors to consult the book [27:80, 22:314, 11:209]. One set of comments suggests that the book is so commonly cited because it is the only book on the subject [11:10].

Norms and best practices are referenced implicitly and explicitly in threads on the forum. Implicit references refer to a right and a wrong way of doing things, often asking the forum to rule on best practice. Examples include:

"Is this the correct practice?" [19:63]

"I want to start naming the files correctly RIGHT from the start." [22:307]

Explicit references directly ask about or refer to best practices and norms:

"Which DVD's have the consensus of being the best for image archiving and storage?" [22:80]

"I am an avid fan of Apple's Aperture. But I have been considering using Photo Mechanic as it is the industry norm for image archiving and such." [27:29]

"whats the best practice regarding importing using LightRoom" [27:42]

Evidence for the existence of a shared ethos in the social world of photography include an accepted authoritative source for guidance, implicit and explicit reference to norms, and evidence of common practices across the community. These common practices are described in more detail in the following section.

5   Paradigm and the personal photo collection

This section uses the Paradigm guidelines for creators of personal collections to organize discussion of common collection management practices described on the forum. Shared common practices indicate

some level of consensus regarding best practice across the dispersed community. The remainder of the section is a structured comparison of the photographers' best practices with (a) practices recommended by digital preservation professionals (i.e. the Paradigm guidelines); and (b) key findings from the existing photowork and PIM literatures. Each subheading is one of the Paradigm guidelines. The second guideline, *Manage your emails*, is omitted because the topic is not discussed on the forum.

5.1 Organise and name files appropriately

These guidelines are concerned with making it easier to find and manage files, as well as with maintaining the context of documents over time. The desire to get organized drives some forum authors to consider or reconsider workflows, choose tools for managing photos, and develop systems for storing, naming, and otherwise describing photos [11:38, 22:84]. The question of why they feel it is important to develop a very organized system is not discussed; it is unclear whether being organized is a strategy for refinding, part of a long-term keeping plan, part of the ethos of serious photography, a general cultural virtue, or a blend of all of these. The result, however, is that attention to organizing collections as they are built makes it easier to maintain the collection in the present and for the long term.

Naming files and folders Part of designing a workflow is deciding on a standard directory structure for organizing photos, a standard method of naming directories in the structure, and a standard way of naming files. Strategies for naming files and naming and structuring directories are discussed in detail on the forum, usually when some author asks for advice on the best strategies [22:110, 22:311, 22:97].

The most common way home mode photographers organize film photos is in their photo processing envelopes (Rodden and Wood 2003; Rose 2003; Vroegindeweij 2001). In the digital environment, the role of the photo envelope is played by the directory created when photos are downloaded. This directory is often automatically created and named with a date, usually the date of download. These directories are the primary organizing method used by home mode digital photographers, with little renaming of individual files (Frohlich et al. 2002; Kirk et al. 2006; Rodden and Wood 2003; Vroegindeweij 2001). This method of organizing arranges images in chronological order, collocating images from specific events, with very low effort on the part of the photographer.

Most descriptions of organization strategies described by forum authors are more complex. Multi-level directory structures are employed by many. It is most common to arrange top-level directories by date; however, the granularity of date used varies. Some have YearMonthDay directories at the top level [11:214, 22:97], while others have Year-Month-Day directory structures [22:309]. Other directories are

24

added for aspects such as event [22:97] (and sub-event [22:309]) and project name [22:296, 22:127]. Outside these repeated date-based structures, some authors report having additional special directories. These are used to organize work (in progress, completed, to print, to burn) [22:278], photo format [22:140], or photo quality (best photos, ok photos, bad photos), and may be forced to the top of the directory list by adding appropriate characters to the beginning of the directory names [22:97, 22:127]. If photos are grouped by type (portrait, landscape) at the top level [22:140], date directory structure is added beneath each category [22:97]. These practices echo previous findings on the importance of date for organizing photos, as well as manipulation of directory order and reuse of directory structures (Jones et al. 2005).

While there are forum authors who do not rename individual image files [22:309], much more individual file renaming is reported on the forum than in the literature on home mode digital photographers. This is likely because forum authors are more technically savvy and/or are using photo processing or organization tools with batch renaming features. As with directory names, it is common to included dates in file names [11:156, 22:311]. One author includes the entire creation timestamp, down to the second, in each file name [11:214]. Other information in file names includes photographer name [11:156, 11:213], project name/number [11:194, 11:214], and media card ID number [11:194]. Some unique sequential number is also commonly part of the file name. This may be retained from the name generated by the camera, or created by the renaming script.

Overall, discussion on the forum implies that serious photographers are interested in, and often meeting, The Guidelines' best practices for file organization and naming. They create meaningful names. One author explains, "I designed this naming scheme so that the file name alone is sufficient to contain the critical data necessary to find the picture" [11:214]. These names have standardized forms. One author reports maintaining a list of codes for use in his file names [11:217]. The date format yyyymmdd seems to be most commonly used and there is some discussion about why it is better than other ways of recording date [11:218]. The main downfall of serious photographers' file naming is that the names and directory structures often become long and/or complex. Using tools with automatic rename capabilities removes some of the difficulties of dealing with this complexity, but it can still be problematic [11:215], as one author attests: "my file names became so long (in Lightroom) that the Seagate would not back-up hundreds of files" [11:126].

Have a plan for handling versions Version proliferation—the existence of multiple versions of the same information, which may also be scattered across locations—is a challenge in personal collections (Kirk et al. 2006; Marshall 2008; Williams et al. 2008) Forum authors create and sometimes keep several

versions of the same file [22:157. Photos may be saved at various stages of processing to allow for rolling back to earlier state [22:150]. Images color corrected for specific uses may be saved as different versions [22:129]. The lower-resolution images emailed to friends or posted online might be saved as further versions in different sizes [22:116].

Though version proliferation is present in serious photographers' collections, it does not seem to be a problem. Versions are the products of a carefully planned workflow, so a system for dealing with them without confusion is in place. Making changes to file names is one method used to organize versions [20:85, 11:66]. Another strategy is to create new directories for new versions when working with batches of images [19:153]. Some combine both of these strategies [11:215].

Make data self-documenting This guideline is essentially concerned with creating and managing metadata in order to maintain information about the context of files. We know that home mode photographers generally do not spend much time on metadata-related activities. They may aspire to annotating and labeling all of their photos or arranging them with captions in albums; however, this level of description is not necessary to support the use of the home mode photo collection, and is thus endlessly deferred until another day (Frohlich et al. 2002; Rodden and Wood 2003; Rose 2003).

Forum authors, on the other hand, express much interest in creating and using metadata to describe their photos. This is demonstrated in several ways. First, there are numerous threads asking about the best way to tag, keyword, or otherwise describe images [11:186, 22:123, 12:131]. Second, authors report spending a lot of time and effort describing their images [22:126, 22:191, 27:63]. Third, tagging and keywording are taken seriously. One author refers to "keywording methodology," [22:133] while another suggests a formal approach and use of a controlled vocabulary [22:288]. Fourth, adding keywords to all photos is regularly mentioned as part of workflow [27:70, 27:74]. In short, adding metadata to images appears to be the norm for a well-managed collection:

> "By mangaed [sic] I mean uniquely named, keyworded and captioned, with my copyright
> and contact data embedded, rated with either stars and color tags (or both) organized into
> folders and collections." [27:78]

While home mode photographers typically do not search their collections (Bentley, Metcalf, and Harboe 2006; Kirk et al. 2006; Rodden and Wood 2003), preferring to rely on browsing by chronology (Miller and Edwards 2007) for images taken recently (Cunningham and Masoodian 2007). Forum authors, on the other hand, seem concerned with increasing the findability of their photos [11:192, 22:303].

They report wanting to search their photos on a number of attributes: camera used to shoot image [11:87, 22:62], camera settings used (f-stop, focal length, shutter speed) [11:143, 22:250], pre-digitization file format [22:115], image resolution [22:114, 27:33], and rating of image [22:293, 22:16]. Previous studies found that serious amateur photographers who post photos on Flickr report tagging photos solely for the benefit of others (Cox, Clough, and Marlow 2008; Miller and Edwards 2007). The amount of time forum authors report spending on image description and their desire to be able to find specific photos in flexible ways highlights the fact that the use of Flickr to manage a sub-set of an image collection tells us little about what people do with their actual image collections.

A commitment to describing photos as part of a photo workflow lands the serious photographer into metadata issues beyond the ken or care of most people managing personal collections. First, four major metadata standards may be used to manage personal image collections: Exif [11:189, 22:225], IPTC [27:69, 22:50], XMP [12:123, 22:112], and ICC [11:222, 13:85]. EXchangeable Image File (Exif) is used to embed information such as camera type, exposure settings, and creation timestamp in a photo when it is captured. International Press Telecommunications Council (IPTC) metadata is an industry standard for describing aspects of photos not described by Exif. Some applications still read and write IPTC metadata directly, but recent implementations express the IPTC core standard information using Extensible Metadata Platform (XMP), an RDF-based standard for embedding content and rights metadata in the documents they describe. International Color Consortium (ICC) profiles describe the color spaces of image input, display, and output devices. They must be managed, but are typically not manipulated by those using them.

Despite the existence of three standards for photo description, elements important to photographers such as tags, ratings, and post-processing data are not standardized, and are implemented differently in different applications. This is problematic for photographers who begin with a basic photo organizing application and later want to upgrade to a more powerful tool. Authors report being unable to transfer keywords/tags between applications [11:32, 22:72] or retain album category information when migrating to a new computer [16:62]. When considering organizational strategies and software choice, serious photographers take into account the portability of their metadata [12:123].

Further, serious photographers must protect the authenticity of their image metadata from applications that overwrite or delete metadata 27:69]. This usually happens when photos are being migrated from one application to another or from one storage location to another. Sometimes undesired metadata changes are made by the operating system based on use of the file. Replacement of creation timestamp with modification or replication time is both common [13:149, 22:100] and particularly problematic because

creation date is an essential piece of information for sorting and finding photos. Further, the operations that make such changes are often run on large batches of photos, causing widespread damage.

Another layer of metadata complexity is that most RAW formats are writable only by their respective cameras. Since they cannot be edited, the authenticity of the original image data is ensured. The downside is that metadata—such as descriptions of versions, post-processing transformations, and image subjects—cannot be embedded in the images. The solution is to write metadata into another file, known as a sidecar, that is linked to the RAW file described. This creates another associated file to manage. Sidecars can become disassociated from the files they describe or otherwise lost [19:211]. Adobe's DNG RAW image format can embed all metadata in one file with the image (or not—it is up to the user) [22:112]. DNG is sometimes referred to the equivalent of the physical film processing envelope, as it contains image and metadata [19:219, 19:220], and this is one of its recognized strengths.

This metadata complexity and the ways in which serious photographers learn about and deal with it have not been discussed in previous work on this group. The photo industry is aware of the importance of image metadata and has formed a working group to explore issues. The group has released guidelines for managing image metadata in the design of image creation and manipulation tools (Group 2009). There is a large body of research on image metadata within LIS.[9] Of particular interest is the 2006 Digital Images Archiving Study (Anderson et al. 2006), targeted to professionals working with digital images in cultural heritage institutions. None of these groups references either of the others. Making links between these communities could improve the outlook for the future of all digital image collections.

Delete what's not important   The Guidelines recommend making choices about what materials to keep for the long term, advice at odds with some views that it is better to keep everything, since the availability of abundant storage makes it cheaper to do so than to spend time making keeping/weeding decisions (Tan et al. 2007). Making decisions about what to keep or delete is difficult (Bruce 2005), complicated by the fact that the value of items in personal collections is nuanced, highly contextual, and changes with time (Marshall 2008). Yet a collection by definition implies selection (Lee 2005), and as difficult as culling is, many people do not want to keep everything (Marshall 2008); as we gain the capability to record everything in perpetuity, it becomes important to highlight the emotional necessity of being allowed to forget (Dodge and Kitchin 2007).

Some form of culling occurs even if the photographers never delete photos. For this reason, there is a distinction between the terms sorting, weeding, and selecting. Weeding is a binary decision: delete or

---

[9]See section 1.4 for a number of references to this work.

keep. Selecting involves choosing some photos for a particular treatment—processing, posting online, printing, long term storage—without taking action on the unselected files. Sorting is simply a process of putting photos into categories, referred to in (Kirk et al. 2006) as "triage." These three are closely related. Sorting may be used just to organize images, or it may be a first step in a selection or weeding process. Weeding is a specific kind of selection: selection for deletion. Conversely, selection for long-term retention could be viewed as a passive weeding decision.

Forum authors report a range of strategies for narrowing down the collection. It is common practice to delete photos of poor technical quality [22:274, 16:204, 16:221]. Some edit "ruthlessly" [16:46]:

> my thinking is to cull out not just ruthlessly but to the point of fanaticsm. if you keep your
> best then this becomes the target to which you have to shoot that quality to in the future. it
> will over time make you a better photographer and not just a shutter snapper. [16:225]

Others keep nearly everything [16:214, 16:220]. Some point out that as time passes, photos once thought unimportant or even bad may increase in sentimental or monetary value [16:224, 16:203]. Others mention the usefulness of even poor images as source material for photo editing [16:210] or for instructional or entertainment value:

> Then there are photos so bad that they're worth keeping. Mine are collected together in a set
> I call When Decisive Moments Go Awry [16:224].

Most authors seem to fall somewhere in the middle. They are nervous about the finality of deletion, but see the value of culling the collection: "I view deletion as a cathartic process that much improves my perception of whats left" [16:217]. Culling is typically an iterative process done in multiple stages during the workflow. One author explains, "What persists in my working archive has to survive at least four to five rounds of cutting and grading" [16:210]. If files are to be deleted, it general recommended practice is to allow time between shooting and deletion. One strategy is to move files to be deleted into a holding directory for a waiting period before deletion [16:207, 22:231, 32:81].

There are a few reports of deleting images from the camera pre-download [16:201, 16:218, 32:80], but some express reservations about making quality judgments based on camera display [32:80]. Instead, most weeding, sorting, and selecting is done post-download. Some photographers acquire laptops (and now are considering netbooks [16:127, 16:100]) to use for initial weeding on shoots or when traveling [14:74, 15:80, 15:155, 19:97]. The laptop functions as a larger camera display so that the quality of images can be confidently assessed [16:127]. It may also be used to move images from memory cards

(that are used in the camera) to (often multiple) external hard drives for safe keeping [16:127, 15:62, 15:86, 22:216]. At this point, poor quality photos (out of focus, etc.) are often deleted [16:201, 19:225].

After initial download, decisions are made about which images to develop from RAW into another format for further editing or other use [13:162, 19:222, 19:223]. Sub-collection creation—for example, choosing some photos to add to Flickr photostream, include in a photobook, or share with a photo club [19:230, —requires sorting. Finally, selection may take place before photos are archived. There are several threads about which file formats to archive (discussed further below).

5.2   3. Select suitable formats, software

This Guideline highlights choice of software and file formats affects the ease of long-term preservation. The literature on home mode personal photo collections does not treat the topic of file formats at all. One study mentioned in passing that a participant had both JPGs and PSDs (Marshall 2008), but overall a digital photo is described as a digital photo as though all digital photos are the same.

The situation is very different in serious photography. Forum authors most frequently mention working with JPG and RAW files. RAW is not itself a format, but a term used to refer to an image format used to encode the raw data from the CCD in a digital camera. NEF (Nikon), CR2 (Canon), and DNG (Adobe) are the most frequently mentioned specific RAW formats. PSD and TIFF are also mentioned a number of times, and a wide array of other formats are mentioned once or twice. Forum authors also mention having film negatives, transparencies, slides, prints from film, gallery prints of digital images, photobooks, and various multimedia objects made using their photos. This underscores a broader need to understand the entire photography-related collection and not only the digital images.

File formats are an important issue on the forum; the topic is its own category, but the topic is discussed all over the forum. Discussion of file formats roughly falls into five types. First, there are questions about the general advantages and disadvantages of different formats [19:53, 22:322, 27:27]. Longevity of the format is not mentioned. Second, there is discussion of which format(s) to use for active work [19:181, 19:183, 26:28]. This topic centers around quality tradeoffs when editing, saving, and printing images.

The third type concerns incompatibility of software and file formats [19:102, 20:47, 22:198, 16:117]. These problems stem mainly from the proprietary nature of native camera RAW formats. For example, Nikon uses the NEF raw format, while Canon has the CR2 raw format. These formats are not equivalent. Some include compression, and others use encryption. New cameras often introduce RAW formats

that are not readable by existing software. These issues can sometimes be solved by downloading free software updates. Sometimes introducing new tools or steps into the workflow is necessary to use a new camera. This quote illustrates the convoluted steps required to open a Canon RAW file in an older version of Photoshop (note that he still can't get the files open):

> I clearly am missing something. I was told that I could convert my .cr2 files to adobe raw and edit them in PS CS. Here's what I did: (1) I have Adobe DNG converter 4.3.1 (2) I have the Camera RAW plugin for CS (3) I convert my .cr2 file to a .dng file (works fine, no problems) (4) I try to open in PS CS ... They will not open in PS CS, I get the error "not the right kind of document" [19:19]

Forum authors generate a number of versions of images in several different formats, and the fourth type of question is about which of these format(s) to retain in backups/archives [19:233]. The final type of question centers on whether or not to convert camera RAW files to DNG files [19:63, 19:78]. This question is related to the previous ones and is discussed further below.

Learning to work with RAW formats is a milestone in the amateur career. Shooting RAW is the digital equivalent of developing your own film and enlarging your own prints in the darkroom, whereas shooting JPG is like dropping your film off at the drug store. RAW is the raw data from the CCD, allowing the photographer to make adjustments to exposure, sharpness, color tone, and more before the data is rendered into an image format, or digital "print." This process does not change the original RAW file in any way. Multiple versions of an image can thus be pulled from a RAW file, as multiple prints can be made from a film negative. The RAW file is the negative of digital photography—the original, authentic representation of the light present at the moment of exposure (Lang 2007).

Serious photographers are understandably attached to their digital negatives, but those digital negatives are inherently problematic from an interoperability (and thus preservation) standpoint. Problems opening current RAW files with commonly used software in a straightforward manner bodes ill for the long-term usability of these files. To some extent, this is recognized in the photography community. Adobe created the Digital NeGative (DNG)— a standard format for storing RAW data. The company also released a free tool to convert closed RAW formats into DNG (DNG Converter). As an open format for RAW data that enables embedding of metadata, DNG could greatly increase the long term safety of digital photographs. However, the question of whether or not DNG has long-term viability as a standard format is still open. While a number of manufacturers of cameras and software have adopted the format in the past year, major players Canon and Nikon still use their proprietary formats. Also, there is some

resistance to the DNG format in the photography community.

Unless all major camera brands begin natively writing DNG, there will be resistance to the format within the photography community. If a camera does not create original images in DNG, transforming its RAW files to DNG is a step away from the authentic negative [20:109]. DNG does not replace the original RAW file in the minds of forum authors. To ensure access to some RAW version of their photos if the camera native format becomes obsolete, some forum authors retain the camera native and DNG files in long term storage [12:162, 20:100]. Others view this as unnecessary because it involves retaining another large file and they are not concerned about not being able to open the original file in the future [12:170]. One author believes he has this problem solved:

> "A quick trick to make sure you will always be able to open your raw file in the future is to get a cheap mac or pc that can run your needed software as today, put it in a plastic box with cd copy of your software, tape it and put it in the attic. If one day you need to acces those raw that are not supported in Vista 28 or in OS 17 call WildCat at this time, just open the box and get your raw from it, that way you will save space, money and hard drive for other use than DNG file..and save major time today not converting all your file" [19:218].

Another writes:

> I also don't understand the rationality of converting raw into dng [. . .] What is the point of making life difficult. Scientifically and logically, the most "efficient" way of doing anything is the one that involves least step(s) without sacrificing the quality and durability of the outcome [19:238].

Unfortunately, this attitude may be directly contributing to the "durability of the outcome." Native RAW formats are far too numerous and proprietary to depend upon for the long term. Current preservation standards for raster images recommend retention of image files in TIFF or JPEG2000 formats. The latter is not mentioned by forum authors. TIFF is mentioned, but relatively little compared to JPG and various RAW formats.

Some forum authors recognize attractive qualities of the TIFF format: it is an open standard [19:245, 20:84, 20:108], offers lossless compression [20:84], and is recommended by Lightroom staff as the best target format to use when converting RAW files [19:246]. Authors do report converting RAW files to TIFF for printing [12:90, 19:181] or editing [19:185, 22:150]. Others convert older non-RAW files or digitized images to TIFF [12:116, 12:124, 14:169, 22:168]. A number of authors report archiving TIFFs

along with their RAW files [19:231, 19:235, 22:278, 36:71]. Sometimes others question their wisdom in doing so [12:169]. Others keeping TIFFs are warned that they should also be keeping their RAW files [22:144]. Among a number of forum members, the attitude toward the safest image format currently available is:

> Tiff is completely pointless to me. [12:170]

> i dont keep any TIF anymore as i can just rerun a action from the PSD to get them. [22:321]

The ethos of the photography world is to honor the sanctity of the original image. And, if you have the negative, you can always make a reprint. The implications of this for long-term preservation of these collections are discouraging.

5.3   Backup your files

This section summarizes common practices described by forum authors using the terms *backup*, *archiving*, and *storage*. Overall, authors make no clear distinctions between the terms (see section 4.7 on page 22), so they are discussed all together here.

People fear losing their data and understand that regular backups and taking other steps to safeguard data for the long term are important for protecting data against loss; however, their actions do not always mirror their values. Many people back up infrequently or not at all (Kaye et al. 2006; Kirk et al. 2006; Marshall 2008), or in ways that don't safeguard files. For example, they may "backup" to a different folder on the same hard drive or store all backups in the same physical location (Williams et al. 2008). As discussed in section 4.5 on page 20, forum authors also fear data loss [12:31, 34:73], but the norm is to take extensive steps to guard against the possibility. This supports the finding in (Williams et al. 2008)that more technically proficient users are more diligent about backing up.

Be selective and organized   Creators of home mode and general personal collections tend toward unsystematic approaches. Reasons include ad hoc backup via email and use of social websites, time and effort required to backup, lack of appropriate storage media, lack of technical knowledge, and unrealistic beliefs about the replaceable of data (Kaye et al. 2006; Marshall 2008).

In contrast, creation of a backup routine is part of the well-planned photography workflow for serious photographers. The level of selectivity for backup purposes is unclear from the forum data. There seems to be no typical practice, as preference for selectivity in choosing which files to retain varies widely across individuals (see section 5.1).

Make copies on portable media   Use of portable media for backup is commonly reported in the literature: floppy disks (Marshall 2008; Williams et al. 2008), zip drives (Marshall 2008), external drives (Marshall 2008; Williams et al. 2008), CDs (Kirk et al. 2006; Marshall 2008; Williams et al. 2008), DVDs (Marshall 2008; Williams et al. 2008), and alternative computers (Marshall 2008; Williams et al. 2008). Floppies are too small for use with today's photos and zip drives are out of date. Use of neither is described on the forum. Also, forum authors appear to be technically competent and engaged enough to migrate files from system to system, so do not collect old computers as external media. Forum authors heavily use external hard drives and, to a lesser extent, optical media.

External hard drives are generally a cheaper, simpler solution for replicating the large image collections of serious photographers, and appear to be the most popular form of portable media for backup on the forum. Optical media are considered by some to be too slow to burn and read, too unreliable, and too small [12:99, 22:176]. Having at least two backup drives is considered the best practice [22:146]. Backup drives need not be portable—some forum authors have multiple internal drives and RAID setups to ensure systematic file replication [12:42, 16:76, 12:133, 13:104]. These systems do not allow for physical off-site storage, however. Two authors mention using several hard drives [16:146, 11:177]. One reports using five drives [13:130], and two use six [22:106, 22:77]. Redundancy is a key criterion in evaluating a backup system on the forum [12:42, 12:150]. Drive reliability is another factor important to forum authors [11:125, 12:99, 16:78]. The true test of any backup system is of course whether it provides an adequate restoration after a loss of data. One forum author did report a mostly successful restore [12:155].

Though external drives are generally considered better for backup, forum authors do back up to DVDs [22:60, 11:105, 12:127], a combination of DVDs and CDs [15:246, 22:141, 22:159], or CDs [19:157]. A common concern is the long term, "archival" reliability of optical media. Authors have a sense that some brands are better than others and ask for recommendations [12:159, 22:214, 22:80]; a favorite brand is MAM-A. Forum authors are also concerned with the effect of labeling optical media on their archival properties [13:113, 12:148]. Finally, optical media are also sometimes used in concert with external hard drives to rehouse replicated data [11:159, 12:144].

Store a copy off-site   The Guidelines suggest storing a backup at some physical remove from other backups. This protects against physical catastrophe. Some creators of general collections do this by keeping important information in safe deposit boxes (Marshall 2008). Forum authors also report this strategy [22:183]. One provides extensive guidelines for a backup system using safe deposit boxes [37:74].

Others move hard drives to off-site storage locations [11:147, 16:146, 22:155].

Another method of off-site storage is backing up to a remote networked location. Typical personal collection backup strategies such as using webmail as an offsite repository or uploading things to web services like Flickr (Marshall 2008; Williams et al. 2008) are clearly unfeasible given the scale of the serious photographer's collection. There is some interest on the forum in large-scale web storage hosting services [11:57, 11:190, 12:140, 22:165], but most backup systems rely mainly on physically accessible hard drives and optical media.

Use data synchronization services or software   Some forum authors are using data synchronization services more advanced than what the Guidelines recommend. Apple's Time Capsule, an automatic backup drive for Macs, is used by several forum authors [22:236, 17:34, 22:31]. Others have automated systems for backing up to an external hard drive [16:146, 22:132] Still others describe their network attached storage (NAS) automated backup systems using tools such as Drobo [11:54, 12:155, 13:130, 22:177]. These strategies were also observed in a minority of advanced users in (Williams et al. 2008).

5.4   Look after your hardware and media

This Guideline stresses the fallibility of hardware and media, and gives suggestions on how to minimize the chances of hardware and media failure. Forum authors cite the inevitable fallibility of hard drives as one of the reasons behind their backup plans [11:223, 11:224, 12:171]. The only routine for monitoring and upgrading hardware mentioned so far is a plan to annually replace the primary internal hard drive used [11:147]. That said, there are surprisingly few reports of data loss due to hardware failure. Forum authors frequently cite running low on space and increasingly sluggish performance as reasons for replacing hardware [12:42, 12:67, 16:172, 14:196]. The failure rate of optical media is also mentioned by forum authors [12:173, 12:74], and they are aware that things like labeling and brand quality can affect the longevity of the media [11:225]. One author recommends checking burned DVDs on a regular basis [22:213] and another stresses the importance of finding out the best writing speed in order to avoid errors in burning optical media [22:215].

5.5   Administer your system

Much of workflow design falls under this guideline. As discussed above, forum authors report careful planning of their systems and updates. There are no reports of viruses or badware in the forum data analyzed.

5.6 Consider using passwords and encryption devices

There is little discussion of dealing with passwords and encryption on the forum. This fits with previous findings that amateurs do not consider their photos to be private documents (Miller and Edwards 2007). Subcollections are frequently made available via web services with password-protected administration access, but the original images are retained in the local collection.

5.7 Be aware of intellectual property rights and privacy

As mentioned above, serious photographers do not consider their photos private and do not report encrypting or password protecting them. In general personal collections, the inclusion of materials created by others requires care to avoid violation of intellectual property laws. The overwhelmingly self-created composition of the serious amateur's collection means this is not a large concern.

Instead, forum authors express concern about about identifying and protecting their photographs as their intellectual property. One asks about the proper time in the photo processing workflow to embed watermarks in photos to be shared with others [11:128]. Others discuss the use of copyright/rights metadata [19:221, 27:78, 22:82]. One author expresses concern that the terms of service of a new photosharing service gives the service too many rights for use of user-posted photographs [29:41]. Finally, some forum authors asked questions about managing and interpreting software licenses [13:145, 15:175, 16:74, 19:58]. While the applications themselves are not part of the collection, it is good practice to maintain some record of software purchase, licensing, and installation details. This enables photographers to make sure they are using the software legally, and also gives them evidence to use if requesting replacement license codes in the case of hardware failure or theft.

5.8 Keep up to date

Amateur photographers' thirst for knowledge and professional standards suggest that they will keep hardware up to date as best they can. High end cameras, image management and organization software, and the computers needed to run them are expensive. Cost is a one of the most frequent criteria mentioned in evaluative statements about tools and services, but amateurs want to have the newest and best they can afford. New hardware can be incompatible with older software. This fact, along with the desire for new features, ensures that software will be kept up to date. Greater danger lurks in the archives, however: old file formats silently become unusable as new applications drop support for them. Keeping file formats fresh by migrating them to newer formats flies in the face of the ethos of photographs. It would be viewed

as compromising the original digital negatives.

5.9   Handling legacy digital files

Some forum authors have already run into issues of dealing with legacy digital files:

> I have a number of vintage 90's Kodak PhotoCD disks and I realize that the time has come
> to process them out of that format for archival storage. I noticed it was time to do something,
> when I went to process an image using Photoshop CS4. I installed the CS3 version of the
> PhotoCD plugin and processed the image I needed, but this isn't going to work forever
> [12:115].

Photo CD is a proprietary format no longer readable by current versions of PhotoShop or Lightroom. The author quoted above and one other [19:79] were advised that they would need to copy the files to another form of media and that they would need to process the files individually in order to transform them into an open format. It seems that such problems will be come more frequent as proprietary RAW files begin to age.

5.10   11. Consult professionals

The Guidelines encourage collection creators to consult preservation professionals for guidance in maintaining their collections. There is no evidence that forum authors consider doing this. The forum is a place for consulting professional photographers, but as discussed above, their advice sometimes conflicts with what preservation professionals would advise. Forum authors exhibit a hunger for information on how to best take care of personal digital collections. They cite *The DAM Book* as a resource and want to know if similar books exist [11:77]. Development of personal photo collection management guidelines addressing the complexities and values of serious photography in the presentation of current preservation best practice could be a way for archival professionals to reach out to this community.

6   Conclusion

This chapter reports partial, preliminary analysis of posted threads from a single category of a single photography forum. Practices described on the forum cannot be assumed to represent the practices of serious photographers in general. The content is shaped by those who speak out, ask questions, and share

advice—a small subset of serious photographers. Thus, no firm conclusions can be drawn from this work at this time.

Revisiting the research goals from the beginning of the chapter, however, shows that the study has been successful. The first goal was to determine if there is anything new and important to learn from studying the personal collection management of amateur photographers. The answer to this is clearly yes. Many practices are described on the forum that have not been described in previous work on personal photo collections. These include proactive collection management extensive description of collections using several metadata standards, successful handling of multiple versions of photos, development of extensive backup systems, and design of entire computing systems to support the tasks of processing, organizing, and storing image collections. In particular, further work should explore whether these practices are reported by photographers who do not share their practices on the forum. Other questions highlighted include:

- What are photographers' long-term goals or plans for their collections? It is clear that maintaining the collections is important to photographers, but the forum data do not tell how the collections are used or what creators expect or hope for their collections in the long-term future

- How do serious photographers learn about and deal with the complexity of the metadata required for managing their personal collections? To what extent do they create, manipulate, and use metadata? How do their image descriptions compare with those seen on photo sharing sites and institutional collections?

- Is there some metaphor or model in the serious photographer's understanding of and approach to backing up and archiving that could be used to design tools and/or educational materials to help home mode photographers better protect their collections from data loss?

- How do serious photographers contrive breakdown situations in order to plan for and prevent future collection management problems?

The second research goal was to learn about the norms and expectations of the social world of photography regarding photography-related collection management. While it is not clear exactly who the forum authors are, a number identify themselves as either professionals or amateurs, indicating the forum is a shared mediated communication space for the social world of photography. Norms and best practices are there to aspire to; whether or not most photographers follow practices recommended on the forum, certain recommended and expected practices were evident. These include: management and maintenance

of an organized, well-described collection; using a RAW workflow; having extensive backup plans; and keeping up with new hardware and software.

In depth discussion of the third research goal—development of a qualitative data analysis template for future work—is beyond the scope of this chapter. It is worth noting, however, that fairly large changes were made to the initial template's concepts of a) situations leading to collection management activities, and b) the scope of activities considered relevant to collection management.

Finally, there appear to be three areas of difficulty for which LIS professionals might be able to help create solutions. First, as one forum author explains [22:289], there are two distinct yet complementary paradigms for organizing image collections. On one hand, there is the vertical, visible, hierarchical file system structure of named directories containing named subdirectories and named files. On the other hand, a horizontal, abstract, flexible structure can be brought to the images using applications that leverage metadata. One problem is that the distinction between and importance of both of these organization methods is rarely explicitly spelled out. New adopters of operating systems and various photo organization and manipulation tools exhibit confusion about exactly how their images are being organized—what the underlying model of the organization system is. This leads to frustration, tedious attempts to force applications to do things they were not designed for, and in the worst cases, data loss. This is an aspect of collection management that applies not just to photos. Explaining these different paradigms, the purposes of each, and what kind of tools manipulate which aspects of photo organization should be incorporated into educational materials for personal collection creators. As experts in organizing information, LIS professionals could contribute to developing such explanations and learning objects.

Second, the image metadata situation highlights the disconnects between the end-user, the photo industry, and cultural heritage image collection communities. Each of these has knowledge and perspective on the issues of organizing images that may be useful to the others, but they do not appear to communicate. LIS professionals could initiate some attempts to bridge these communication gaps, or, at the very least, can incorporate the knowledge of the other two in its own understanding of the problem space.

Finally, the serious photography community's privileging of the RAW format and disregard for preservation-friendly formats is troubling for those of us concerned with the long-term viability of these collections. Photographers will not be convinced to abandon archiving RAW files as digital negatives, nor should anyone try to convince them: to not keep the negative would fly in the face of the ethos of serious photography. What preservation professionals could do here is strongly encourage the practice of keeping a high quality TIFF version of an image in addition to the negative. In addition, we could become involved in and advocate for a standard RAW format so that we can be confident of our ability

39

to preserve what is considered the most authentic version of an image.

This study also demonstrates once again that interesting new findings that run counter to LIS' "common" knowledge can be found by looking at serious leisure pursuits. Though the study is far from conclusive, it raises a number of new questions and suggests avenues for practical action aimed at helping to ensure the long-term safety of personal image collections.

References

Adams, Anne, Sally Jo Cunningham, and Masood Masoodian (2007). *Sharing, privacy and trust issues for photo collections*. Working paper No. 01/2007. Hamilton, New Zealand: University of Waikato, Department of Computer Science. URL: http://hdl.handle.net/10289/59. See pp. 7, 15.

Adobe (2008). *TechNote: Optimize performance of Photoshop CS3 on Windows XP and Vista*. Adobe TechNote kb401088. Accessed 25 March 2009. Apr. 2008. URL: http://kb.adobe.com/selfservice/viewContent.do?externalId=kb401088. See p. 17.

Anderson, Sheila et al. (2006). *Digital images archiving study*. Tech. rep. AHDS: Arts and humanities data service, p. 163. URL: http://ahds.ac.uk/about/projects/archiving-studies/index.htm. See pp. 7, 16, 28.

Angus, Emma, Mike Thelwall, and David Stuart (2008). "General patterns of tag usage among university groups in Flickr". In: *Online information review* 32.1, pp. 89–101. See p. 7.

Apple (2009). *iPhoto '09*. Software application. Accessed 21 March 2009. URL: http://www.apple.com/ilife/iphoto/. See p. 7.

Baca, Murtha (2002). *Introduction to art image access: Issues, tools, standards, strategies*. Los Angeles: Getty Research Institute. See p. 7.

Barreau, Deborah K. (1995). "Context as a factor in personal information management systems". In: *Journal of the American Society for Information Science* 46.5, pp. 327–339. ISSN: 0002-8231. See p. 4.

Beagrie, Neil (2005). "Plenty of Room at the Bottom? Personal Digital Libraries and Collections". In: *D-Lib Magazine* 11.6, unpaged. URL: http://www.dlib.org/dlib/june05/beagrie/06beagrie.html. See p. 17.

Bentley, Frank, Crysta Metcalf, and Gunnar Harboe (2006). "Personal vs. commercial content: the similarities between consumer use of photos and music". In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. New York: ACM Press, pp. 667–676. See p. 26.

Boardman, Richard and M. Angela Sasse (2004). ""Stuff goes into the computer and doesn't come out": A cross-tool study of personal information management". In: *Proceedings of CHI 2004*, pp. 583–590. See p. 4.

Bruce, Harry (2005). "Personal, Anticipated Information Need". In: *Information Research* 10.3. See p. 28.

Burgess, Jean E. (2007). "Vernacular creativity and new media". PhD thesis. Creative Industries Faculty, Queensland University of Technology, p. 288. URL: http://eprints.qut.edu.au/archive/00010076/. See pp. 5, 9, 17.

Chalfen, Richard (1987). *Snapshot Versions of Life*. Bowling Green, OH: Bowling Green State University Popular Press. See p. 4.

Chen, Hsin Liang and Edie M. Rasmussen (1999). "Intellectual access to images". In: *Library trends* 48.2, pp. 291–302. ISSN: 0024-2594. See p. 7.

Cox, Andrew M., Paul D. Clough, and Jennifer Marlow (2008). "Flickr: A first look at user behaviour in the context of photography as serious leisure". In: *Information Research* 13.1, unpaged. URL: http://informationr.net/ir/13-1/paper336.html. See pp. 6, 7, 17, 27.

Cunningham, Sally Jo and Masood Masoodian (2007). "Identifying personal photo digital library features". In: *Proceedings of the 2007 conference on Digital libraries*. ACM, pp. 400–401. See p. 26.

Cutrell, Edward, Susan T. Dumais, and Jaime Teevan (2006). "Searching to eliminate personal information management". In: *Communications of the ACM* 49.1, pp. 58–64. See p. 4.

Davies, Julia (2006). "Affinities and beyond! Developing ways of seeing in online spaces". In: *E-learning* 3.2, pp. 217–234. See pp. 5, 9.

Dodge, Martin and Rob Kitchin (2007). "'Outlines of a world coming into existence': pervasive computing and the ethics of forgetting". In: *Environment and Planning B-Planning & Design* 34.3, pp. 431–445. ISSN: 0265-8135. URL: http://www.envplan.com/abstract.cgi?id=b32041t. See p. 28.

Enser, Peter G. B. et al. (2007). "Facing the reality of semantic image retrieval". In: *Journal of Documentation* 63.4, pp. 465–481. See p. 7.

Frohlich, David and Jacqueline Fennell (2007). "Sound, Paper and Memorabilia: Resources for a Simpler Digital Photography". In: *Personal and Ubiquitous Computing* 11.2, pp. 107–116. ISSN: 1617-4909. See p. 7.

Frohlich, David et al. (2002). "Collaborating around collections: Requirements for photoware". In: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. New York: ACM Press, pp. 166–175. See pp. 4, 15, 24, 26.

Gelber, Steven M. (1999). *Hobbies: Leisure and the Culture of Work in America*. New York: Columbia University Press. 374 pp. See p. 4.

Giddens, Anthony (1984). *The constitution of society: Outline of the theory of structuration*. Berkeley: University of California Press. 402 pp. See p. 19.

Google (2009). *Picasa 3*. Software application. Accessed 21 March 2009. URL: http://picasa.google.com/. See p. 7.

Griffin, Michael Scott (1987). "Amateur photography and pictorial aesthetics: Influences of organization and industry on cultural production". PhD thesis. University of Pennsylvania. See p. 5.

Grinter, Rebecca E. (2005). "Words about Images: Coordinating Community in Amateur Photography". In: *Computer Supported Cooperative Work* 14.2, pp. 161–188. See p. 5.

Group, Metadata Working (2009). *Guidelines for Handling Image Metadata*. Accessed 21 March 2009. URL: http://www.metadataworkinggroup.org/pdf/mwg_guidance.pdf. See p. 28.

Hammer, Jessica (2007). "Agency and authority in role-playing "texts"". In: *A new literacies sampler*. Ed. by Michele Knobel and Colin Lankshear. New literacies and digital epistemologies. New York: Peter Lang, pp. 67–93. See p. 4.

Hart, Chris, Michael Schoolbred, and David Butcher (1998). "The bibliographical structure of fan information". In: *64th IFLA General Conference*. IFLA. URL: http://www.ifla.org/IV/ifla64/020-123e.htm. See p. 4.

Hartel, Jenna (2007). "Information activities, resources & spaces in the hobby of gourmet cooking". PhD thesis. University of California, Los Angeles, p. 273. See p. 4.

Heidegger, Martin (1978). *Being and time*. Oxford: Basil Blackwell. See p. 20.

House, Nancy A. Van (2007). "Flickr and public image-sharing: distant closeness and photo exhibition". In: *CHI '07 extended abstracts on Human factors in computing systems*. ACM, pp. 2717–2722. See p. 6.

Jones, Ian and Graham Symon (2001). "Lifelong Learning as Serious Leisure: Policy, Practice and Potential". In: *Leisure Studies* 20.4, pp. 269–283. ISSN: 0261-4367. See p. 3.

Jones, William P. (2008). *Keeping Found Things Found: The Study and Practice of Personal Information Management*. Amsterdam: Morgan Kaufmann. 430 pp. See pp. 3, 4, 20.

Jones, William P. et al. (2005). "Don't take my folders away!: organizing personal information to get things done". In: *CHI '05 extended abstracts on Human factors in computing system*. New York: ACM Press, pp. 1505 –1508. See p. 25.

Jorgensen, Corinne L. (1999). "Retrieving the Unretrievable: Art, Aesthetics, and Emotion in Image Retrieval Systems". In: *IS&T/SPIE Internet Imaging IV*. Ed. by Bernice Ellen Rogowitz and Thrasyvoulous Pappas. Bellingham, WA: SPIE, pp. 348–355. See p. 7.

— (2003). *Image Retrieval: Theory and Research*. Lanham, MD: Scarecrow Press. 340 pp. See p. 7.

Kari, Jarkko and Jenna Hartel (2007). "Information and higher things in life: Addressing the pleasurable and the profound in information science". In: *Journal of the American Society for Information Science and Technology* 58.8, pp. 1131–1147. See p. 3.

Kaye, Joseph 'Jofish' et al. (2006). "To have and to hold: Exploring the personal archive". In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. Ed. by Rebecca E. Grinter et al. New York: ACM Press, pp. 275–284. See pp. 4, 15, 33.

King, Nigel (1998). "Template analysis". In: *Qualitative methods and analysis in organizational research*. Ed. by Gillian Symon and Catherine Cassell. London: Sage. See p. 9.

Kirk, David et al. (2006). "Understanding photowork". In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. New York: ACM Press, pp. 761–770. See pp. 4, 7, 15, 16, 24–26, 29, 33, 34.

Kotro, Tanja (2007). "User Orientation Through Experience: A Study of Hobbyist Knowing in Product Development". In: *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments* 3.2, pp. 154–166. See p. 3.

Krause, Michael G. (1988). "Intellectual problems of indexing picture collections". In: *Audiovisual librarian* 14.2, pp. 73–81. See p. 7.

Krogh, Peter (2006). *The DAM Book: Digital Asset Management for Photographers*. Sebastopol, CA: O'Reilly. See pp. 16, 23.

Kustritz, Anne (2003). "Slashing the Romance Narrative". In: *The Journal of American Culture* 26.3, pp. 371–384. See p. 4.

Lambert, Ronald D. (1996). "Doing family history". In: *Families* 35, pp. 11–25. See p. 4.

Lang, Karl (2007). *Rendering the Print: the Art of Photography*. Technical paper. Adobe. URL: http://wwwimages.adobe.com/www.adobe.com/products/photoshop/family/prophotographer/pdfs/pscs3_renderprint.pdf. See p. 31.

Layne, Sara Shatford (1994). "Some issues in the indexing of images". In: *Journal of the American Society for Information Science* 45.8, pp. 583–588. ISSN: 0002-8231. See p. 7.

Lee, Hur-Li (2005). "The concept of collection from the user's perspective". In: *Library quarterly* 75.1, pp. 67–85. ISSN: 0024-2519. See p. 28.

Marcum, James W. (2002). "Beyond visual culture: The challenge of visual ecology". In: *Portal-Libraries and the Academy* 2.2, pp. 189–206. ISSN: 1531-2542. See p. 8.

Marshall, Catherine C. (2008). "Rethinking personal digital archiving, part 1". In: *D-Lib Magazine* 14.3/4, unpaged. URL: http://www.dlib.org/dlib/march08/marshall/03marshall-pt1.html. See pp. 16, 17, 21, 22, 25, 28, 30, 33–35.

Marshall, Catherine C., Sara Bly, and Francoise Brun-Cottan (2006). "The long term fate of our personal digital belongings: Toward a service model for personal archives". In: *Proceedings of IS&T Archiving 2006*. Springfield, VA: Society for Imaging Science and Technology, pp. 25–30. See p. 21.

McKechnie, Lynne E. F. et al. (2002). "Research Method Trends in Human Information Literature". In: *The New Review of Information Behavior Research: Studies of Information Seeking in Context*. Vol. 3. Proceedings of ISIC IV, The Fourth International Conference on Research in Information Needs, Seeking and Use in Different Contexts, Lisbon, Portugal, 11-13 September, 2002. Taylor Graham, pp. 113–126. See p. 3.

McKemmish, Sue (1996). "Evidence of me". In: *Archives and Manuscripts* 24.1, pp. 28–45. ISSN: 0157-6895. See p. 15.

Miller, Andrew D. and W. Keith Edwards (2007). "Give and Take: A Study of Consumer Photo-Sharing Culture and Practice". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, pp. 347–356. See pp. 4, 6, 21, 22, 26, 27, 36.

Murray, Susan (2008). "Digital images, photo-sharing, and our shifting notions of everyday aesthetics". In: *Journal of Visual Culture* 7.2, pp. 147–163. See p. 6.

Neal, Diane R. (2006). "News photography image retrieval practices: Locus of control in two contexts". Unpublished doctoral dissertation. Denton: University of North Texas. See p. 7.

Nightingale, Virginia (2007). "The cameraphone and online image sharing". In: *Continuum: Journal of media & cultural studies* 21.2, pp. 289–301. See p. 7.

Nikšić, Hrvoje (2007). *GNU Wget*. Software package. Accessed 30 January 2009. URL: http://www.gnu.org/software/wget/. See p. 11.

Paradigm (2008). *Workbook on digital private papers*. Tech. rep. Oxford: Oxford University. URL: http://www.paradigm.ac.uk/workbook/index.html (visited on 05/27/2009). See pp. 16, 18.

Rafferty, Pauline and Rob Hidderley (2007). "Flickr and Democratic Indexing: Dialogic Approaches to Indexing". In: *Aslib Proceedings* 59.4/5, pp. 397–410. See p. 7.

Redmiles, David et al. (2005). "What Ideal End Users Teach Us About Collaborative Software". In: *GROUP'05: Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, pp. 260–263. See p. 3.

Rodden, Kerry and Kenneth R. Wood (2003). "How Do People Manage Their Digital Photographs?" In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, pp. 409–416. See pp. 7, 24, 26.

Root, Josh (2008). *Terms and Conditions of Use*. Accessed 30 January 2009. photo.net. URL: http://photo.net/info/terms-of-use. See p. 11.

Rose, Gillian (2003). "Family photographs and domestic spacings: a case study". In: *Transactions of the Institute of British Geographers* 28.1, pp. 5–18. ISSN: 0020-2754. See pp. 24, 26.

Rosenblum, Barbara (1978). *Photographers at Work: A Sociology of Photographic Styles*. New York: Holmes & Meier. 144 pp. See pp. 5, 6.

Schwartz, Dona B. and Michael Scott Griffin (1987). "Amateur Photography: The Organizational Maintenance of an Aesthetic Code". In: *Natural Audiences: Qualitative Research of Media Uses and Effects*. Ed. by T. Lindlof. Communication and Information Science. Norwood, NJ: Ablex, pp. 198–224. See p. 5.

Slater, Don (1991). "Consuming Kodak". In: *Family Snaps: The Meaning of Domestic Photography*. Ed. by Jo Spence and Patricia Holland. London: Virago, pp. 49–59. See pp. 5, 6.

Stebbins, Robert A. (1982). "Serious Leisure: A Conceptual Statement". In: *Pacific Sociological Review* 25.2, pp. 251–272. ISSN: 0030-8919. See p. 3.

— (1998). *After Work: The Search for an Optimal Leisure Lifestyle*. Calgary: Detselig Enterprises. 184 pp. See p. 3.

— (2007). *Serious Leisure: A Perspective for Our Time*. New Brunswick, NJ: Transaction Publishers. See pp. 3, 5, 6, 17, 22.

Tan, Desney et al. (2007). "Save everything: supporting human memory with a personal digital lifetime store". In: *Personal Information Management*. Ed. by William P. Jones and Jaime Teevan. Seattle, WA: University of Washington Press, pp. 90–107. See p. 28.

Teper, Jennifer Hain (2004). "Newspaper Photo Morgues: A Survey of Institutional Holdings and Practices". In: *Library Collections, Acquisitions, and Technical Services* 28.1, pp. 106–125. ISSN: 1464-9055. See p. 7.

Thomas, Angela (2007). "Blurring and Breaking Through the Boundaries of Narrative, Literacy, and Identity in Adolescent Fan Fiction". In: *A New Literacies Sampler*. Ed. by Michele Knobel and Colin Lankshear. New Literacies and Digital Epistemologies. New York: Peter Lang, pp. 137–165. See p. 4.

Torrey, Cristen et al. (2007). "How-To Pages: Informal Systems of Expertise Sharing". In: *ECSCW 2007: Proceedings of the Tenth European Conference on Computer Supported Cooperative Work*. Ed. by Liamand Wagner Bannon et al. London: Springer, pp. 391–410. See pp. 3, 4.

Trace, Ciaran B. (2008). "Resistance and the Underlife: Informal Written Literacies and Their Relationship to Human Information Behavior". In: *Journal of the American Society for Information Science and Technology* 59.10, pp. 1540–1554. See p. 4.

Unruh, David R. (1980). "The Nature of Social Worlds". In: *Pacific Sociological Review* 23.3, pp. 271–296. See p. 8.

Urban, Richard (2007). "Second Life, Serious Leisure, and LIS". In: *Bulletin of the American Society for Information Science and Technology*, pp. 38–40. See p. 3.

Van Dijck, J. (2005). "From shoebox to performative agent: the computer as personal memory machine". In: *New Media & Society* 7.3, pp. 311–332. ISSN: 1461-4448. See p. 15.

Vroegindeweij, Sander (2001). *My pictures: informal image collections*. Technical Report HPL-2002-72R1. Palo Alto, CA: HP Labs, p. 105. URL: http://www.hpl.hp.com/techreports/2002/HPL-2002-72R1.pdf. See p. 24.

Williams, Pete et al. (2008). "Digital lives: Report of interviews with the creators of personal digital collections". In: *Ariadne* 55, unpaged. URL: http://www.ariadne.ac.uk/issue55/williams-et-al/. See pp. 15, 21, 22, 25, 33–35.

Winograd, Terry and Fernando Flores (1986). *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, NJ: Ablex. See p. 20.

Yakel, Elizabeth (2004). "Seeking Information, Seeking Connections, Seeking Meaning: Genealogists and Family Historians". In: *Information Research* 10.1, unpaged. ISSN: 1368-1613. URL: http://informationr.net/ir/10-1/paper205.html. See p. 4.

Yoder, D. G. (1997). "A Model for Commodity Intensive Serious Leisure". In: *Journal of Leisure Research* 29.4, pp. 407–429. ISSN: 0022-2216. See p. 5.