

# Best Practices for Digitizing Photographs: A Network Analysis of Influences

Paul Conway; University of Michigan; Ann Arbor, Michigan/USA

## Abstract

*This paper presents the first significant findings from a major study of how the terms and conditions for building collections of digitized visual resources, particularly historical photographs, influence how users judge archival quality, integrity, and value (and hence usefulness) of the digitized resources. The paper articulates criteria for distinguishing best practice documents from general opinion pieces or reports on experiments. The paper then assesses the cumulative influence and power of seventeen guidelines documents on the development of best practice recommendations and a network analysis of the interconnected patterns of influence by individual experts and leading organizations that have sponsored or contributed to the development of community best practices.*

## Understanding Guidelines

In the span of two decades, the keepers of cultural heritage have transitioned their approach to the digitization of visual resources, particularly historical photographs, from rarified experiments to nearly ubiquitous activity. In the United States the Research Libraries Group, Cornell University, the Library of Congress and other large organizations led early experiments. Today, digitization activities are widely dispersed, yet still relatively small in scale and generally uncoordinated. Erway and Schaffner suggest that it is only a matter of time before the scaling up of special collections digitization efforts forces a reassessment of procedures and policies. [22] Lynch asserts, however, that digitization is largely a solved problem. "The research questions are less about how to do it at all and more about how to optimize – how to do it more efficiently or effectively." [28] In the vast number of smaller and more isolated non-profit organizations, significant uncertainty about practices and procedures remains deeply rooted, even as larger organizations position themselves for sustained programs.

Standards governing most of the underlying technological aspects of building and sharing digital collections are in place and maintained by a variety of organizations, including the International Standards Organization [25] and its national counterparts, the World Wide Web Consortium, [40] and the Association for Information and Image Management. [18] Formal standards vary in specificity and their adoption runs the gamut from optional to required, depending on circumstances and the propensity of local practice to reflect the larger context of standards. Chapman suggests that a special challenge for the cultural heritage community is not the lack of standards but the plethora of competing standards. [19] Rada concludes a sequence of case studies on standards development by arguing for a close investigation of how individuals or organizations influence formal processes. [35]

Rather than codify their digital product requirements in national and international standards, libraries, archives, historical centers, museums, and similar organizations have opted instead to develop best practices derived from practical experience. When communicated across institutions through the direct sharing of experience or promulgated more formally through published guidelines, best practices are a time-sensitive community consensus on technical comparability. [38] In an environment of rapid applications development, some see community consensus as superior to formal standards-generating processes, which may suffer from process formalities, questionable openness, and a lack of agility in response to technological changes. [39] Best practices for digital imaging tend to emphasize a combination of procedural recommendations for digitization and a range of recommendations on the significant properties of the image product itself, including but by no means limited to bitmap characteristics, and metadata elements describing digital objects. [30]

Understanding the value and relevance of digitization guidance is becoming more important as cultural heritage organizations position themselves to increase the scale and variety of retrospective digitization programs to encompass visually or technically complex resources such as photographs, maps, and audiovisual resources. The United States has no national coordinating agency comparable to the UK's Joint Information Systems Committee, no natural forum for assessing cross-institutional digitization practices in widely dispersed organizations, and no mechanism for reviewing ongoing digitization practices comparable to initiatives supported by Europe's DELOS. [19] The absence of a research agenda that encompasses the creation of digitized products based on cultural heritage resources may also be problematical. [36]

A comprehensive literature search yielded no evidence that any research has ever been undertaken to assess the value or impact of digitization guidelines and best practices produced over the past fifteen years in the United States. Liu's assessment of digitization guidelines is haphazardly documented and reaches no specific conclusions about the recommendations within the guidelines. [26] Lopatin assembles a selective literature focusing on project management, funding, selection for digitization, metadata, and related matters but draws no conclusions about the origins, development or application of quality guidelines. [27] Puglia and Rhodes speculate that guidelines documents show a trend from low to high special resolution, from low bit to high bit image scanning, and from scanning for a specific purpose to a more "use neutral" manner. Their conclusion demonstrates the need for systematic investigation of past practices and future directions. "It is a little humbling to look back and admit that we

are still asking many of the difficult questions that we were asking over a decade ago.” [34]

This paper reports the findings of one part of a larger study that seeks to understand the relationship between the ways that archivists and librarians build collections of digitized historical resources and the judgments that users make about the quality and archival integrity of these resources in digital form. The research reported here is limited to identifying a set of published guidelines specifying digitization parameters for historical photographs and identifying the individual and organizational influences on the development of these documents. This initial focus is motivated by nearly ubiquitous activity over a fifteen year period and the wide availability of specific guidance on best practices for photographic digitization. A forthcoming article will present a methodology for normalizing digitization parameters and assessing change in recommended digitization practices over time. Cumulatively, the deconstruction of digitization guidelines establishes a foundation for evaluating the impact of digitization workflow on the creation of collections of digitization surrogates that end-users must themselves evaluate for relevance and usefulness after retrieval from an image digital library.

### Research Methodologies

The project involved three discrete steps: (1) identify target guideline documents by applying selection criteria to a large group of candidate documents; (2) create an analysis database from normalized information extracted from the target set; and (3) extract a set of social networks from the data set.

#### Identifying Target Guideline Documents

Broad and iterative surveys of the existing literature in photographic digitization processes and projects in the cultural heritage sector generated a list of 95 potential guideline documents. Search focused on resources identifiable through open web search engines, indexed in Wilson Library Literature and ISI Web of Science, compiled on preservation oriented web-based web portals such as Preserving Access to Digital Information (PADI) and Conservation Online (CoOL), [31] or cited or acknowledged in published articles. [29] The search excluded promotional and marketing documents produced by commercial enterprises, particularly from vendors of digitization services and the purveyors of digitization software, scanner hardware, and integrated workflow systems. This study makes no explicit claim for search comprehensiveness.

The search results set, ranging in publication from 1992 to 2007, includes a heterogeneous mix of formal reports, published articles and newsletter stories to white papers and sometimes marginally attributed web pages with no physical analog. The identification of guidelines resulted from the application of three criteria to the search results set:

- (1) *Promulgation*. Guidelines must be released to the public in some combination of print and electronic versions. Excluded are documents restricted to a narrow community through password protection, blog posts, and wiki entries.
- (2) *Declaration*. Guidelines declare themselves directly as such or are cited by multiple sources as a guideline or statement of best practices. News articles describing guidelines or best practices documented elsewhere and the findings of small

pilot projects whose results are neither intended to be generalized nor cited as so in other sources are excluded.

- (3) *Specification*. Guidelines include specific (occasionally multi-tiered) recommendations on digitizing original photographic prints or negatives. A document must specify, at minimum, resolution and bit depth parameters for prints or positive or negative photographic film. Excluded are guidelines that focused solely on digitizing microfilm or books.

The three filter criteria yield seventeen target documents, dating from 1995 to 2006. Two of the seventeen are available only in hard copy (one is out of print); the remaining fifteen are freely available through the web. Table 1 lists the seventeen documents in sequential (chronological) order, showing the short name used in some of the illustrations and the date of publication. Full bibliographic citations are included in the references. [1-17] Since completing the analysis an additional document that meets the criteria has been published. [39]

**Table 1: Photograph digitization guidelines**

Short Name	Ref	Pub Date	Sequence
RLG 95	1	1995	1
FREY/REILLY 96	2	1996	2.3
CORNELL 96	3	1996	2.5
LoC RFP 96	4	1996	2.8
COLUMBIA 97	5	1997	3
NARA 98	6	1998	4.4
FLEISCHHAUER 98	7	1998	4.8
FREY/REILLY 99	8	1999	5
KENNEY/RIEGER 00	9	2000	6
UIUC 01	10	2001	7.4
CDL 01	11	2001	7.8
NINCH 02	12	2002	8
WESTERN STATES 03	13	2003	9.4
IFLA 03	14	2003	9.8
NARA 04	15	2004	10
CDL 05	16	2005	11
LoC 06	17	2006	12

#### Creating an Analysis Database

With the support of a graduate student assistant, the investigator combed each of the seventeen documents (and associated web pages) for evidence of individuals named as authors of guidelines, authors of cited references related to photograph digitization, project consultants, project participants, or other contributors. For each named person, a code recorded his or her institutional affiliation at the time of the publication of the report and the role played in the publication. Subunits within a given organization (e.g., Image Permanence Institute) were coded by parent institution (e.g., Rochester) and home state. Each document also was coded based on the source of the recommendations. A guideline qualifies as “empirical” if recommended practices are based on pilot studies or other experiments conducted by the authors or other named participants and organizations. A guideline is “synthetic” if its recommendations represent a synthesis of recommendations by

other persons or organizations. A "hybrid" guideline contains a combination of empirically derived recommendations and synthesis of other work.

Data from each of the seventeen target documents was double coded, cross verified, and imported into a MySQL database using PERL scripts to extract data from text-based coding sheets. Figure 1 illustrates the simple structure of the analysis database. The completed database contains 341 records with 494 data values. The seventeen guidelines documents yielded 98 discrete individuals and 54 discrete organizational names, encompassing the affiliations of named individuals, sponsoring and funding organizations, and publishers.

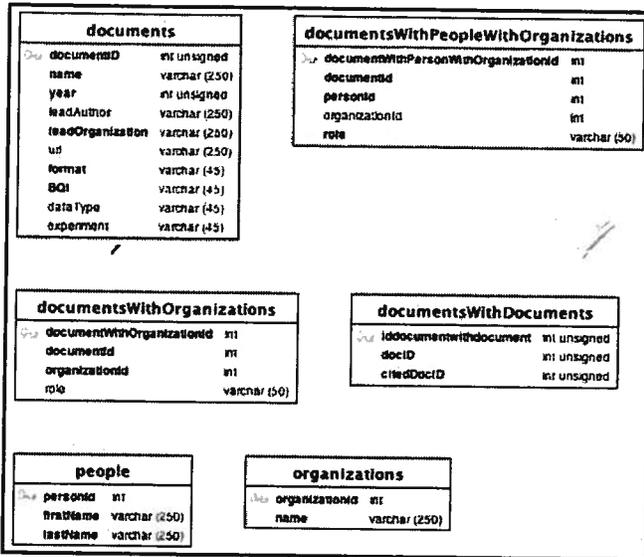


Figure 1. Data model

### Social Network Analysis

The network visualization utilizes GUESS and Pajek open source software toolkits. [33] Pajek is primarily designed for large network analysis, but includes a large number of features that

support exploratory analysis. [21] GUESS is an exploratory data analysis and visualization programmed in a variation of Python. [24] Network analysis approaches include: (1) the PageRank algorithm developed by Google to weight the importance of a web page based on the number of links to that page; [32] (2) Girvan and Newman's collaboration model for co-authorship networks in science; [23] and (3) measures of betweenness, or the directness of the connections between individuals, organizations, and their geographic locations.

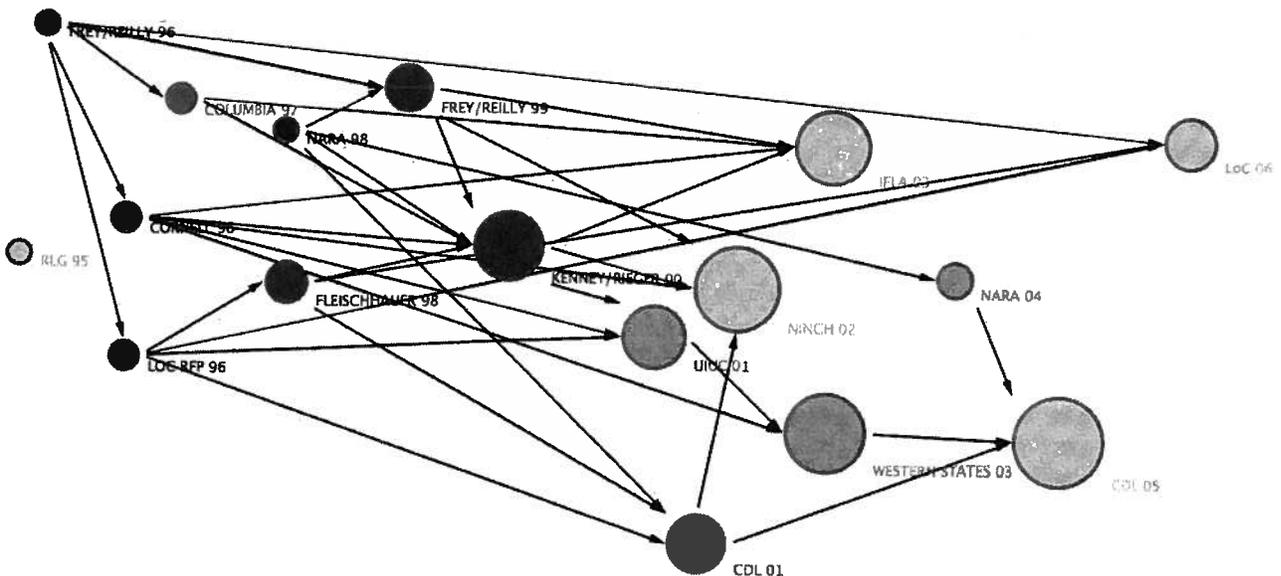
## Findings

### Cumulative influence of guidelines

The first exploratory analysis illustrates the extent to which published guidelines are influential over time. The hypotheses are that older documents influence newer ones and that overall influences converge toward a single document. The stronger evidence for this hypothesis the stronger the claim for a unified community consensus drawn from cumulative experience.

Figure 2 (below) is a visual representation of the cumulative influence of the seventeen guidelines over a twelve year period. The figure shows the patterns of citations of older documents by more recent documents. Circles are sized based on the cumulative number of citations to past documents (out-degree). Documents represented by larger circles (e.g., CDL 05, NINCH 02) have cited more of the seventeen documents explicitly (as direct footnotes or references) or implicitly (by citing intermediate-aged documents that themselves cite older documents). The arrows connect the older documents to the newer ones, showing both the number of citations (i.e., number of arrow sat each circle) and the source of the citations. For example, the document labeled Kenney/Rieger 00 cites four earlier documents and is itself cited by two older documents (LOC 06, NINCH 02), where the influence chain ends.

The shade of the circles code the relative influence of the individual documents overall. Dark is most influential, light is least influential. The figure shows that some older guidelines documents are more influential than others. RLG 95, the oldest document in



the study, is an outlier not cited by any other document. The seminal Frey/Reilly 96 guideline demonstrates the most direct and indirect influence over the formulation of best practices.

The analysis does not confirm a convergence of recommendations to a single document, but rather multiple and multifaceted recommendations oriented toward different audiences. A number of specific conclusions emerge from a graphic representation of influences. First, two documents (IFLA 03 and NINCH 02) seem to reflect the greatest influence from earlier documents; both of these organizations made an attempt in their guidelines to state overtly the source of their recommendations and to cast their documents as a synthesis of past work. Second, Library of Congress guidelines are the most focused on a chain of internal documents (LOC RFP 96) or consultants (Frey/Reilly 96). Third, guidelines from the California Digital Library (CDL 05) factor in the greatest range of prior work, forming a synthesis that incorporates work done in both university and federal government settings.

### **Influential people and organizations**

The underlying data derived for the project identifies a small number of highly influential individuals and organizations that played a significant role in the development of the target set of guidelines. Seventy-three of the 98 discrete individuals are acknowledged in only one of the seventeen target guidelines. Of the 54 separate organizations named in the documents, 39 are credited only a single time in the set of seventeen target guidelines.

Table 2 lists the five most frequently acknowledged organizations, along with the number of citations in the seventeen guidelines documents. Cornell, Harvard, and the Rochester Institute of Technology together are cited 32 times; the Library of Congress and the National Archives and Records Administration, both federal government agencies are credited as either sponsoring organizations or publishers in a total of 16 citations.

**Table 2: Most influential organizations**

Rank	Name	Citations
1	Cornell	14
2	Harvard	10
3	NARA	9
4	Rochester	8
5	Lib. of Congress	7

Table 3 lists the six most influential individuals, along with their institutional affiliations and the number of times they are cited or credited by name in the seventeen target documents. Stephen Chapman, initially active at Cornell but employed in the Harvard University's Wiseman Preservation Center since 1997, tops the list of people who either authored or consulted on guidelines or prepared specific publications on photograph digitization that were cited in one or more guidelines. James Reilly and Franziska Frey, both of the Image Permanence Institute at Rochester Institute of Technology played important roles as both authors and consultants. Anne Kenney is based in the preservation department of the Cornell University Library, while Steven Puglia is an imaging specialist at the National Archives. Among the most influential imaging experts named in the study, only Michael Ester

of Luna Imaging in California, is not affiliated with a not for profit organization.

**Table 3: Most influential people**

Rank	Name	Affiliation	Citations
1	Stephen Chapman	Cornell/Harvard	7
2	James Reilly	Rochester	4
2	Anne R. Kenney	Cornell	4
3	Steven Puglia	NARA	3
3	Franziska Frey	Rochester	3
3	Michael Ester	Luna Imaging	3

Within this leading group the role of formal preservation programs stands out. Chapman, Kenney, and Puglia are all associated with library preservation programs. The Image Permanence Institute has served as a preservation research and development program since its inception. Before founding Luna Imaging, Ester was an imaging specialist at the Getty, whose mission and programs encompass the preservation of cultural heritage.

### **Expertise network**

The second exploratory analysis exposes networks of individual expertise and the interconnectedness of involved individuals. The social network analysis technique utilized is "actor betweenness centrality" following a method developed by Girvan and Newman. [23] The approach counts all of the minimum paths between two individuals that pass through a third "actor." The visualization tool plots these three-way relationships, emphasizing the most prominent individuals by constructing network nodes and sizing them according to their betweenness within the network as a whole. For a given set of network actors, the more closely associated they are, the tighter the network of expertise. A hypothesis for this exploration proposes that given the overlap between a small number of expert/leaders and their associated support organizations (outlined in tables 2 and 3 above), there exists a very tight and circumscribed network for guidelines development.

Figure 3 charts the results of the analysis. The six most prominent expert/leaders listed in Table 3 are labeled in the figure with their associated nodes. The network of people involved in the development of digitization guidelines over a twelve year period is quite tight and consists of approximately eight sub-networks that are connected directly through one or more prominent expert/leaders. Stephen Chapman, the most frequently referenced individual in the study, for example, is a collaborator with two groups of people who never collaborated with anyone outside of their group other than him. Chapman's betweenness score increases further because of his direct relationship with other expert/leaders who themselves are significant connectors with other groups.

Five of the six most prominent expert/leaders functioned actively in the 1990s as peers and colleagues, leveraging and reinforcing each other's expertise as they worked on individual or collaborative projects involving many other individuals. Michael Ester emerges as an expert/leader through his seminal publications on photograph digitization that served as critical touchstone

documents for a number of the guidelines documents. The network analysis portrayed in Figure 3 is not sufficiently fine-grained to reflect the disparate roles of author, leader, or collaborator played by individuals in the study. A limit in this exploration is posed by the guidelines documents themselves. If a given person's contribution to the document is not acknowledged directly by the authors, then they do not appear in the analysis and their perhaps subtle influence is lost.

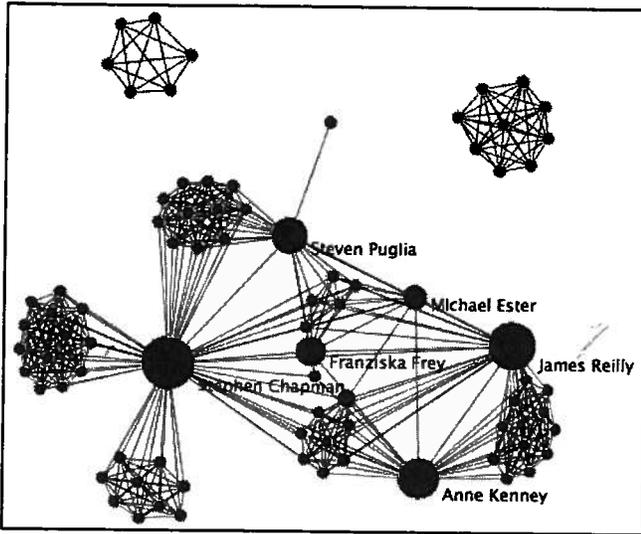


Figure 3. Six individual experts are prominently identified in the guidelines

The 73 individuals in the study who do not emerge as expert/leaders in themselves cluster into at least eight separate sub-networks; six of these sub-networks are connected indirectly to each other through the most prominent individuals in the study. The two collaborative groups that are not directly connected to the expert/leader network are the guidelines projects at Columbia

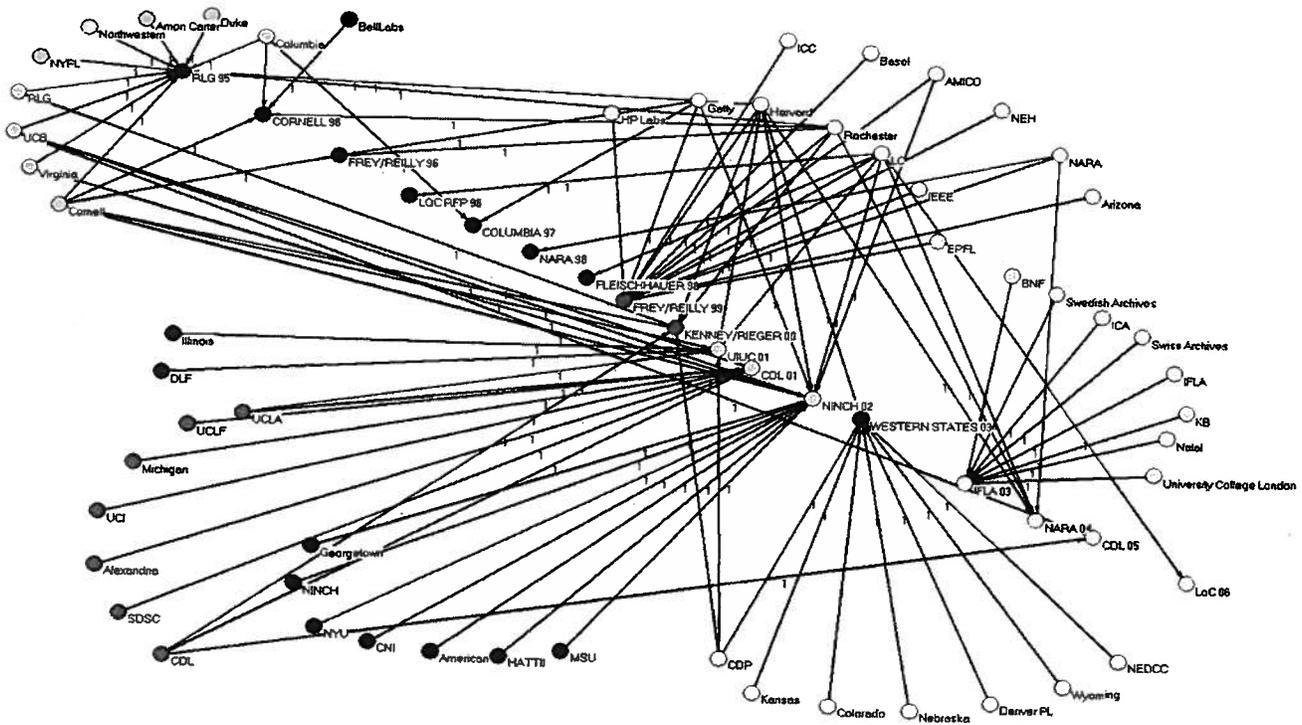
(1997) and IFLA (2003). In both documents, the authors thanked contributing organizations but did not specify individual contributors beyond the working groups that assembled and authored the document. This particular aspect of the data is partially addressed by examining more closely patterns of organizational support for guidelines development.

### Increasing organizational involvement

The third exploratory analysis investigates the extent to which organizational support and buy-in for digitization guidelines expands over time. In the absence of a formal standards development and maintenance program, which centralizes activity in a third party, best practices might be expected to develop iteratively as technology advances and experience increases within a community. A hypothesis for this analysis proposes that over time an increasing number of organizations will participate in guidelines development.

Figure 4 illustrates the organizational involvement in supporting the seventeen documents, either by sponsoring the work of the individuals involved, publishing the results, or hosting the activities behind the guidelines. In the figure, the guidelines documents are aligned in a chronological sequence from left to right. The organizations associated with the development are linked to each document with which they are associated. The graphic shows how clusters of organizations form around particular documents and how some particularly influential organizations are connected to a sequence of guidelines over time. For example, Cornell University was involved in formulating the first guideline in the study (RLG 95) and is also connected to the development of four later guidelines, all of which are synthesis documents

Figure 4. Expanded organizational involvement over time



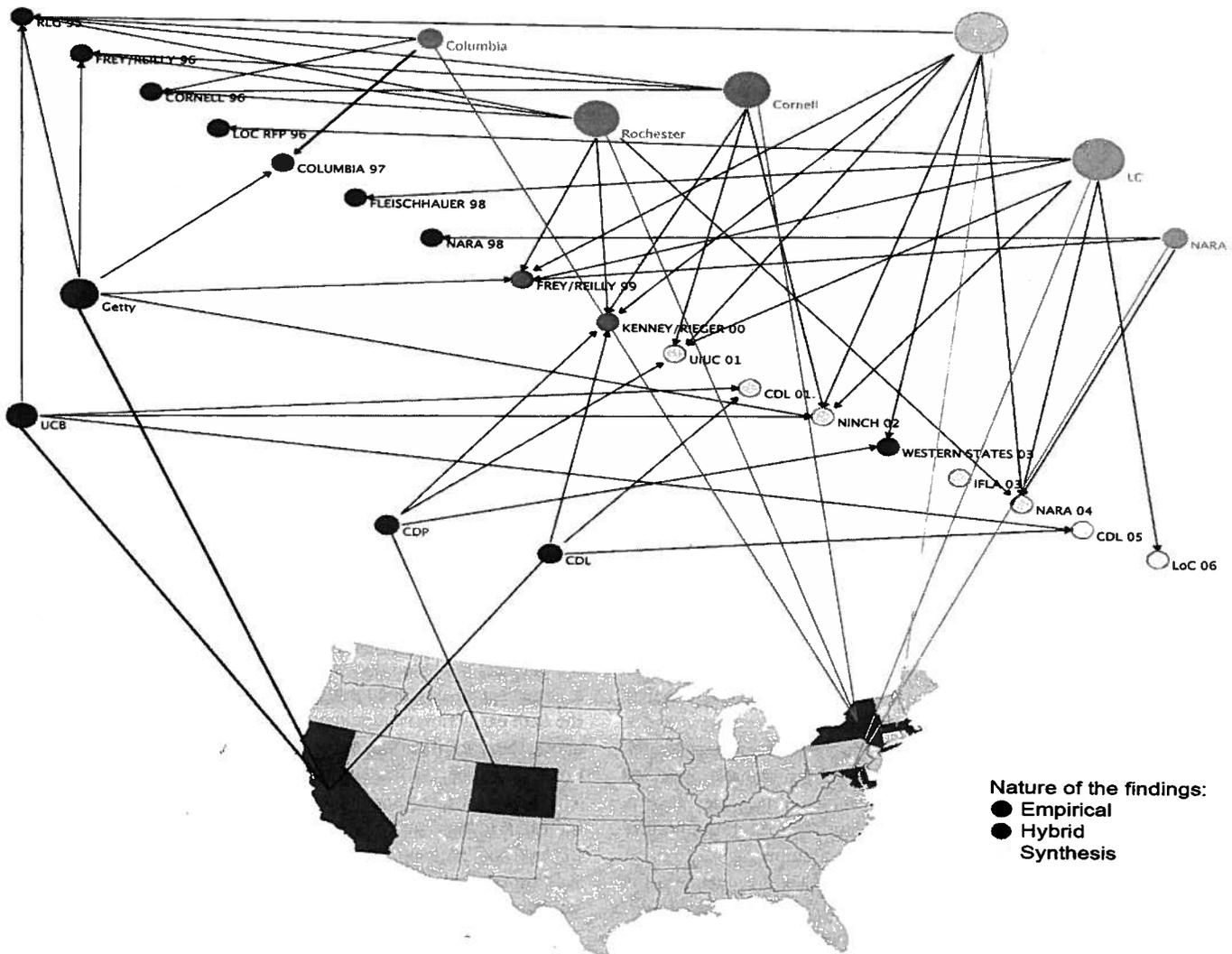
The beginnings of formal guidelines development involved a multi-organizational collaboration hosted by the Research Libraries Group. Four of the members of the 1995 R&D efforts are connected to guidelines developed up to nine years later. Until around 1999, a relatively small number of organizations participated in or sponsored guidelines development. Figure 4 illustrates a very significant expansion in the number and variety of organizational involvement from 2000 onward. With one exception, every guidelines project includes at least one organization with connections to one or more other guideline document. This demonstrates a weak but continuous chain of influence over time. The guideline developed by the International Federation of Library Associations (IFLA) is the only one of seventeen in the study that appears to be an organizationally discrete undertaking. This work of synthesis was developed by an international team of European organizations that worked nearly exclusively with secondary research and without the active

is a derived work with little influence within the US cultural heritage community.

### Geography and the shift to synthesis

A fourth exploratory study started as an effort to investigate patterns of geographic influence. Organizations with three or more citations in the study database are represented in Figure 5. Arrows point both to the documents in which they are cited but also to the geographic states in where they are headquartered. The circle for each organization is sized according to the number of citations in the database. Additionally, each of the seventeen documents is shaded to reflect the nature of the data underlying the individual recommendations; black for fully empirical, light gray for a work of synthesis, medium gray for a hybrid mix of empirical and synthesis.

Figure 5. Guidelines synthesis has a bi-coastal flavor



engagement of consultants or expert/leaders from the United States. The strong interconnections in the IFLA document illustrated in Figure 2 and the lack of interpersonal networking illustrated in Figure 3 add evidence to the conclusion that IFLA 03

The visualization shows an unambiguous trend from guidelines based on experimental research to guidelines that are a strong synthesis of prior work. Two documents (Frey/Reilly 99 and Kenney/Rieger 00) are a rich mix of empiricism and synthesis.

Both documents are extensive collaborations in which the work of many people is synthesized and then supported by direct experimentation by the authors of the guidelines. Later in the sequence, the Western States 03 guideline bucked the trend toward synthesis by structuring a project in which the collaborating organizations conducted specific experiments and then shared the results.

Among the other patterns that emerge from this illustration is the influence of Harvard University in supporting the development of synthesis guidelines. All but one arrow points toward documents developed after 1999. The role of the California Digital Library in supporting the development of synthetic guidelines is also prominent. In the case of CDL, the organization itself serves a coordinating role for the eleven campus California system, so the focus on synthesis is mission-specific. The patterns represented in this figure are determined in part by the institutional affiliations of the top experts identified in the study and listed above in Table 3. The six most influential expert/leaders in the study map tightly on the organizations that demonstrate the greatest involvement in guidelines development.

## Discussion

The most important implication of the study for best practices development is the value of empirical experimentation by a small social network of highly motivated entrepreneur individuals supported by employing institutions. Six important imaging experts, each with a different perspective on the challenge of photographic digitization and supported by three universities and two federal agencies, established the foundation for best practices in a very short period of time. Between 1995 and 1999, these individuals and organizations led experiments and published results that served as a catalyst for a movement toward synthesis. Empirical experimentation was not undertaken in isolation, even if the work was rarely coordinated closely.

With the exception of a single multi-state experimental project, nearly all guidelines development after the year 2000 was focused on synthesizing cumulative experience in the field. Consensus is not a crescendo at a summit meeting or marked by a single publication. Rather, the key to synthesis is an ever-expanding roster of people and organizations willing to participate in collaborative guidelines development or willing to endorse such activity. Expanding participation is evidence of acceptance of a common approach, even if the specifications of the approach are neither transparent nor unitary in nature. The findings suggest that consensus on best practices for photographic digitization is not a singular phenomenon, but rather ones that cluster geographically and by type of leading organization. Federal government agencies appear to take one route to synthesis while universities take a separate route. The net result may be a significant dilution of effort and mixed messages to the community of practitioners.

The development of a synthesis on best practices for photographic digitization is inefficient, time consuming, expensive, and inconclusive. The net result of fifteen years of experimentation and synthesis is multiple recommendations, endorsed by a variety of types of organizations. The process of multi-faceted synthesis has no validating mechanisms and no obvious avenue for dissemination and promotion of use. The cultural heritage community identifies and adopts given best

practice recommendations principally by channeling the reputation of the sponsoring organization and the persuasive powers of the principal individuals involved in their development. These limiting factors in part help to explain why archivists, librarians, and curators in smaller organizations continue to assert the inadequacy of guidelines and best practices for digitization while larger organizations implement ongoing programs tailored to local needs.

The major implication of the lack of a widespread community consensus on photographic digitization is the (as yet untested) potential for greater variance in the quality and integrity of the digital products produced by individual cultural heritage organizations than is warranted even by the wide variance in photographic source materials themselves. This heterogeneity of image quality and intellectual description is increasingly apparent in the results of data harvesting projects, such as OAIster [37] that cumulate similar types of resources from multiple organizational sources. Consistency is rarely obtainable, even within the confines of a single large project, when best practices are diverse and subject to local adjustment. Whether heterogeneity poses problems for end users or decreases the preservation value of the digital products are open questions.

A second major implication of the absence of digitization quality and process standards is continuing investment in small scale experimental projects by small and mid-size organizations that either are unaware of or do not necessarily trust the recommendations of existing best practices guidelines. A corollary implication is the impact of seemingly subtle differences in guidelines promulgated by separate large organizations (e.g., California Digital Library and National Archives and Records Administration) on the exchange of digitized holdings across organizations or the aggregation of holdings in third party union catalogs. In the absence of a large scale impact assessment among the users of digitized collections, we can only speculate on whether the inconsistencies across existing best practices create barriers to access and use.

## Further research

This exploratory research is a first step in a more ambitious and extensive effort to assess the impact on end users of the existence of large and growing bodies of digitized cultural heritage resources. Further research will refine the methodologies outlined here. Feedback from readers on the methodologies utilized to select target guideline documents may result in the identification of a larger set of documents for analysis. Similarly, a replication of the data extraction routine may challenge some of the assumptions about the identification of cited names and organizations in the documents and associated web sites of sponsoring organizations. The form of citation analysis utilized in this study is a labor intensive manual process that does not yet lend itself to the forms of analysis that are possible using formal citation databases. Additionally, the "grey literature" nature of many of the target guidelines documents creates an informality to the analysis that may skew the results.

Given the small size of the sample and the heterogeneity of the data extracted from guidelines documents, finding statistical significance in the patterns of citation and influence is not straightforward. A future study would benefit from a larger data set

and from a research design that is tailored explicitly to support statistical analysis of visualizations.

The purpose of the research reported in this study was limited to identifying target documents, specifying a viable analytical technique, and exposing patterns or networks of influence among the people and organizations named in the documents. The study has yet to report on the specific recommendations in the guidelines regarding image quality. Such reporting on variation in digitization parameters (resolution, bit depth, etc.) and in trends over time will help validate the significance of the documents and add important perspective on the implications of inconclusive consensus.

Most important, perhaps, the research to date does not yet speak to the extent to which cultural heritage organizations have used the seventeen documents to guide the digitization of photographic resources. A guideline is only as valuable as the extent to which its recommendations are adopted beyond the organization that make them. Further research will seek to correlate the adoption of particular best practice guidelines across a large and varied set of digital collections in the United States. Anecdotal evidence suggests a very high level of adoption of external guidelines at the project level accompanied by very inconsistent application at the item level. If supported, such findings would lend weight to an argument that the specification of best practices is an inadequate or even inappropriate strategy for ensuring the quality of digitized photographs. If the quality of digitized photographic resources is a definable property that has value to end users, then a common standard for obtaining quality results consistently is a pressing need for the cultural heritage community.

### Photograph Digitization Guidelines

- [1] RLG Digital Image Access Project: Proceedings from an RLG Symposium Held March 31 and April 1, 1995. P. A. McClung (ed.). Mountain View, CA: Research Libraries Group, 1995. [RLG 95]
- [2] J. M. Reilly and F. Frey, Recommendations for the Evaluation of Digital Images Produced from Photographic, Microphotographic, and Various Paper Formats. Washington, D.C: Library of Congress National Digital Library Project, 1996.  
<http://lcweb2.loc.gov/ammem/ipirpt.html> [Frey/Reilly 96]
- [3] A. R. Kenney and S. Chapman, Digital Imaging for Libraries and Archives. Ithaca: Cornell University Library, 1996.  
<http://www.library.cornell.edu/preservation/dila.html> [Cornell 96]
- [4] Library of Congress, National Digital Library Program. Building Digital Collections: A Technical Overview. Request for Proposal RFP 97-9: Conversion of Pictorial Materials. 1996.  
<http://memory.loc.gov/ammem/about/techIn.html> [LOC RFP 96]
- [5] Columbia University Libraries, Technical Recommendations for Digital Imaging Projects. Prepared by the Image Quality Working Group of ArchivesCom, a joint Libraries/AcIS committee. 1997  
<http://www.columbia.edu/acis/dl/imagespec.html> [Columbia 97]
- [6] S. Puglia and B. Roginski, Guidelines for Digitizing Archival Materials for Electronic Access. Washington, DC: National Archives and Records Administration, 1998.  
<http://www.archives.gov/research/arc/digitizing-archival-materials.html> [NARA 98]
- [7] C. Fleischhauer, Digital Formats for Content Reproductions. Washington, DC: Library of Congress, 1998.  
<http://memory.loc.gov/ammem/formatold.html> [Fleischhauer 98]

- [8] F. Frey and J. Reilly, Digital Imaging for Photographic Collections: Foundations for Technical Standards, Rochester: Image Permanence Institute, 1999.  
[http://www.imagepermanenceinstitute.org/sub\\_pages/digibook.pdf](http://www.imagepermanenceinstitute.org/sub_pages/digibook.pdf) [Frey/Reilly 99]
- [9] A. R. Kenney and O. Y. Rieger, (eds.) Moving Theory Into Practice: Digital Imaging for Libraries and Archives. Mountain View, CA: Research Libraries Group, 2000.[Kenney/Rieger 00]
- [10] University of Illinois at Urbana-Champaign, Guidelines for Digital Imaging Projects. December 2001.  
<http://images.library.uiuc.edu/resources/digitalguidev3.pdf> [UIUC 01]
- [11] California Digital Library, CDL Guidelines for Digital Images, July 2001. <http://chnm.gmu.edu/digitalhistory/links/pdf/chapter3/3.29b.pdf> [CDL 01]
- [12] National Initiative for a Networked Cultural Heritage (NINCH) and Humanities Advanced Technology and Information Institute (HATII), University of Glasgow, NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials. Washington, DC: NINCH, 2002.  
<http://www.nyu.edu/its/humanities/ninchguide/> [NINCH 02]
- [13] Collaborative Digitization Project, Western States Digital Standards Group. Western States Digital Imaging Best Practices. Version 1.0, 2003.  
[http://www.cdpheritage.org/digital/scanning/documents/WSDIBP\\_v1.pdf](http://www.cdpheritage.org/digital/scanning/documents/WSDIBP_v1.pdf) [Western States 03]
- [14] International Federation of Library Associations, Guidelines for Digitization Projects for Collections and Holdings in the Public Domain, Particularly Those Held by Libraries and Archives. March 2002. <http://www.ifla.org/VII/s19/pubs/digit-guide.pdf> [IFLA 03]
- [15] S. Puglia, J. Reed, and E. Rhodes, Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images. Washington, DC: National Archives and Records Administration, 2004.  
<http://www.archives.gov/research/arc/digitizing-archival-materials.html> Print copy available from the Digital Library Federation. [NARA 04]
- [16] California Digital Library, CDL Guidelines for Digital Images, November 2005.  
[http://www.cdlib.org/inside/diglib/guidelines/bpgimages/cdl\\_gdi\\_v2.pdf](http://www.cdlib.org/inside/diglib/guidelines/bpgimages/cdl_gdi_v2.pdf) [CDL 05]
- [17] Library of Congress Technical Standards for Digital Conversion of Text and Graphic Materials, Dec 2006.  
<http://memory.loc.gov/ammem/about/techIn.html> [LOC 06]

### References

- [18] AIIM Standards Program.  
<http://www.aiim.org/standards.asp?ID=28639>
- [19] S. Chapman, Techniques for Creating Sustainable Digital Collections, Library Technology Reports. Chicago: American Library Association, 2004.
- [20] C. Chen, et al., "Digital Imagery for Significant Cultural and Historical Materials," International Journal on Digital Libraries 5 (2005): 275-286.
- [21] W. de Nooy, A. Mrvar, and V. Batagelj, Exploratory Social Network Analysis with Pajek. Cambridge: Cambridge University Press, 2005.
- [22] R. Erway and J. Schaffner, Shifting Gears: Gearing Up to Get Into the Flow. Dublin, OH: OCLC Programs and Research, 2007.

- [23] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks." *PNAS* 99 (12) (June 11, 2002): 7821-7826.
- [24] GUESS: The Graph Exploration System.  
<http://graphexploration.cond.org/>
- [25] International Standards Organization, Standards Development.  
[http://www.iso.org/iso/standards\\_development.htm](http://www.iso.org/iso/standards_development.htm)
- [26] Y. Q. Liu, "Best practices, standards and techniques for digitizing library materials – a snapshot of library digitization practices in the USA." *Online Information Review* 28 (5) (2004): 338-345.
- [27] L. Lopatin, "Library digitization projects, issues and guidelines – a survey of the literature." *Library Hi Tech* 24 (2) (2006): 273-289.
- [28] C. Lynch, "Digital collections, digital libraries and the digitization of cultural heritage information." *First Monday* 7 (5), May 2002.  
[http://firstmonday.org/issues/issue7\\_5/lynch/index.html](http://firstmonday.org/issues/issue7_5/lynch/index.html)
- [29] For example, see: MinervaEC, "Digitisation guidelines – a selected list." Last updated 2004-04-20. Accessed 2008-03-27.  
<http://www.minervaeurope.org/guidelines.htm>
- [30] NISO Framework Advisory Group, *A Framework of Guidance for Building Good Digital Collections*. 2nd edition. Bethesda, MD: National Information Standards Organization, 2004. Available from:  
<http://www.niso.org/framework/framework2.html>
- [31] PADI: Preserving Access to Digital Information.  
<http://www.nla.gov.au/padi/>; CoOL: Conservation Online.  
<http://palimpsest.stanford.edu/>
- [32] L. Page, B. Sergey, M. Rajeev, and W. Terry, "The PageRank Citation Ranking: Bringing Order to the Web." (1998).  
<http://dbpubs.stanford.edu:8090/pub/1999-66>
- [33] Pajek. <http://vlado.fmf.uni-lj.si/pub/networks/pajek/> and <http://pajek.imfm.si/doku.php>
- [34] S. Puglia and E. Rhodes, "Digital imaging – how far have we come and what still needs to be done?" *RLG DigiNews* 11 (1) 15 April 2007.  
<http://www.oclc.org/programs/publications/newsletters/diginews.htm>
- [35] R. Rada, et al., "IT standards development and consensus – three case studies." *StandardView* 2 (1) (March 2004): 50-54.
- [36] S. Ross. "Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries," Keynote Address at the 11th European Conference on Digital Libraries ECDL, Budapest (17 September 2007). <http://www.ecdl2007.org/keynotes.php>
- [37] University of Michigan. OAIster. <http://www.oaister.org/>
- [38] J. West. "The Economic Realities of Open Standards: Black, White, and Many Shades of Gray," in S. Greenstein and V. Stango, *Standards and Public Policy*. Cambridge: Cambridge University Press, 2007.
- [39] K. Wisser.(ed.) *Guidelines for digitization*. Revised edition. North Carolina ECHO. 2007 <http://www.ncecho.org/guide/toc.html>
- [40] World Wide Web Consortium. <http://www.w3.org/>

## Author Biography

*Paul Conway is associate professor of information at the University of Michigan. He has made contributions to the literature on archival users and use, preservation management, and digital imaging. His research interests include representing and interpreting visual and textual resources in digital form, particularly in the context of emerging interdisciplinary scholarship in the humanities. He holds a PhD from the University of Michigan and is a Fellow of the Society of American Archivists.*

# Archiving<sup>2008</sup>

June 24–27, 2008  
Bern, Switzerland

## Final Program and Proceedings

General Chair:  
Rudolf Gschwind, University of Basel



**imaging.org**

Sponsored by the  
*Society for Imaging Science and Technology*

*In cooperation with*

**AIC** American Institute for Conservation of Historic & Artistic Works  
**ALA ALCTS** Association for Library Collections and Technical Services  
**CNI** Coalition for Networked Information  
**DLF** Digital Library Federation  
**DPC** Digital Preservation Coalition

**ECPA** European Commission on Preservation and Access  
**ISCC** Inter-Society Color Council  
**IOP** Institute of Physics  
**MCN** Museum Computer Network  
**OCLC** Online Computer Library Center  
**RPS** Royal Photographic Society