

# Data Management Bibliography

Compiled by Meriel Patrick, [Sudamih Project](#), Oxford University Computing Services

Data Management – General and Policy Issues.....	1
Data Sharing.....	2
Digital Curation and Preservation.....	3
Repositories .....	6
Personal Information Management.....	7
Metadata and Related Issues.....	11

## Data Management – General and Policy Issues

### **“Data Practices, Policy, and Rewards in the Information Era Demand a New Paradigm.”**

Barton, C., R. Smith, and R. Weaver. *Data Science Journal* 9 (2010).

[http://www.jstage.jst.go.jp/article/dsj/9/0/9\\_IGY95/article](http://www.jstage.jst.go.jp/article/dsj/9/0/9_IGY95/article).

Outlines steps that can be taken to encourage responsible data management and open access practices, motivate people to develop systems and services for the common good, and develop professional infrastructure for those engaged in informatics.

### **“The Third DCC-RIN Research Data Management Forum.”**

Donnelly, Martin. *International Journal of Digital Curation* 4, no. 1 (June 29, 2009).

<http://www.ijdc.net/index.php/ijdc/article/view/109>.

A report on the third meeting of the Research Data Management Forum which was held in Manchester, UK on April 30 and May 1, 2009, with an overarching theme entitled "Value and Benefits".

### **“The Data Audit Framework: A First Step in the Data Management Challenge.”**

Jones, Sarah, Alexander Ball, and Cuna Ekmekcioglu. *International Journal of Digital Curation* 3, no. 2 (December 2, 2008). <http://www.ijdc.net/index.php/ijdc/article/view/91>.

Shares experiences of implementing the Framework, reports back on the kind of data issues researchers commonly face, and indicates how the Framework will be further developed before being released for widespread adoption.

### **“A Policy Checklist for Enabling Persistence of Identifiers.”**

Nicholas, Nick, Nigel Ward, and Kerry Blinco. *D-Lib Magazine* 15, no. 1 (1, 2009).

<http://www.dlib.org/dlib/january09/nicholas/01nicholas.html>.

A major finding of the Persistent Identifier Linking Infrastructure (PILIN) project was that policy is far more important in guaranteeing persistence of identifiers than technology. This article outlines a six-point checklist for integrating identifiers into information management.

### **“Skilling Up to Do Data: Whose Role, Whose Responsibility, Whose Career?”**

Pryor, Graham, and Martin Donnelly. *International Journal of Digital Curation* 4, no. 2 (October 15, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/126>.

This paper revisits previous investigations into the roles and responsibilities required by a “data workforce”, presents a representative spectrum of informed opinion from the DCC Research Data Management Forum, and makes some recommendations.

### **“Research funders' policies for the management of information outputs,”**

Rightscom. 2007. <http://www.rin.ac.uk/our-work/research-funding-policy-and-guidance/research-funders-policies-management-information-outpu>.

A report on the policy and practice of major UK research funders in managing the information outputs produced by their supported researchers - on the Research Information Network website.

## **Data Sharing**

### **“Designing for Discovery and Re-Use: the ‘ANDS Data Sharing Verbs’ Approach to Service Decomposition.”**

Burton, Adrian, and Andrew Treloar. *International Journal of Digital Curation* 4, no. 3 (December 7, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/133>.

Describes how the Australian National Data Services (ANDS) is designing systems to support data sharing and re-use, structured in terms of the ANDS Data Sharing Verbs: Create, Store, Describe, Identify, Register, Discover, Access and Exploit.

### **“The Publication of Research Data: Researcher Attitudes and Behaviour.”**

Griffiths, Aaron. *International Journal of Digital Curation* 4, no. 1 (June 29, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/101>.

Discusses the 2008 RIN report, To Share or Not to Share: Publication and Quality Assurance of Research Data Outputs (June 2008) and related issues.

### **“Uncovering User Perceptions of Research Activity Data.”**

Loureiro-Koechlin, Cecilia. *Ariadne* 62 (January 2010). <http://www.ariadne.ac.uk/issue62/loureiroKoechlin/>.

Discusses the outcomes and lessons learned from user tests performed on the Oxford Blue Pages, a tool designed to display information about researchers and their activities at the University of Oxford.

### **“To share or not to share: research data outputs.”**

Research Information Network. 2008. <http://www.rin.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs>.

This report presents the findings from a study of whether or not researchers make their research data available to others, and the issues they encounter when doing so.

## Digital Curation and Preservation

### **“Digital Curation and Preservation Bibliography.”**

Bailey, Charles W., Jr. May 17, 2010. <http://digital-scholarship.org/dcpb/dcpb.htm>.

This bibliography presents selected English-language articles, books, and technical reports that are useful in understanding digital curation and preservation.

### **“Report from the Digital Curation Curriculum Symposium (DigCCurr) 2009.”**

Ball, Alexander, and Michael Day. *International Journal of Digital Curation* 4, no. 1 (June 29, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/108>.

The Symposium theme was "Digital Curation Practice, Promise and Prospects". It dealt with issues from the cutting edge of digital curation research, showcased digital curation tools, and considered how to equip the new generation of information professionals with the necessary skills.

### **“Review of the State of the Art of the Digital Curation of Research Data.”**

Ball, Alexander. June 2, 2010. <http://opus.bath.ac.uk/19022/>.

Looks at digital curation of research data in the context of the data lifecycle, and specifically in the context of data repositories. Part of the JISC ERIM (Engineering Research Information Management) Project.

### **“Research Data Preservation and Access: The Views of Researchers.”**

Beagrie, Neil, Robert Beagrie, and Ian Rowlands. *Ariadne* 60 (July 2009).

<http://www.ariadne.ac.uk/issue60/beagrie-et-al/>.

Presents findings from a UKRDS survey of researchers' views on and practices for preservation and dissemination of research data in four UK universities (Bristol, Leeds, Leicester, and Oxford) and place them in the wider UK and international context.

### **“Keeping Research Data Safe (Phase 2),”**

Beagrie, Neil, Brian Lavoie, and Matthew Woollard. April 2010.

<http://www.jisc.ac.uk/publications/reports/2010/keepingresearchdatasafe2.aspx#downloads>.

A report presenting the results of a survey of available cost information, validation and further development of the Keeping Research Data Safe activity cost model, and a new taxonomy to help assess benefits alongside costs.

### **“Digital Preservation Planning: Principles, Examples and the Future with Planets.”**

Boyle, Frances, and Jane Humphreys. *Ariadne* 57 (October 2008).

<http://www.ariadne.ac.uk/issue57/dpc-planets-rpt/>.

A report on the one-day workshop on digital preservation planning jointly organised by the Digital Preservation Coalition (DPC) and Planets, held at the British Library, on 29 July 2008.

### **“DCC&U: An Extended Digital Curation Lifecycle Model.”**

Constantopoulos, Panos, Costis Dallas, Ion Androutsopoulos, Stravos Angelis, Antonios Deligiannakis, Dimitris Gavriliis, Yannis Kotidis, and Christos Papatheodorou. *International Journal of Digital Curation* 4, no. 1 (June 29, 2009).

<http://www.ijdc.net/index.php/ijdc/article/view/100>.

Reviews two recent models for digital curation introduced by the Digital Curation Centre (DCC)

and the Digital Curation Unit (DCU) of the Athena Research Centre, and proposes a fusion of the two models.

**“Report from the DigCCurr 2007 International Symposium on Digital Curation, Chapel Hill, NC, April 18-20, 2007.”**

Day, Michael. *International Journal of Digital Curation* 2, no. 1 (December 2, 2008).

<http://www.ijdc.net/index.php/ijdc/article/view/28>.

This event was organised as part of the project "Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum,".

**“Digitisation Programme Digital Preservation Study.”**

Digital Preservation Coalition. April 2009. <http://www.dpconline.org/advocacy/digitisation-programme-digital-preservation-study.html>.

A scaled down version of the report on the DPC’s extensive analysis of the projects funded through the JISC Digitisation Programme, including assessment of their plans for access beyond the existence of their project grants.

**“Preservation of Web Resources: Making a Start.”**

Emmott, Stephen. *Ariadne* 56 (July 2008). <http://www.ariadne.ac.uk/issue56/jisc-powr-rpt/>.

Reports on a one-day workshop aimed at all those interested in issues relating to institutional Web resource preservation.

**“A Metadata Approach to Preservation of Digital Resources: The University of North Texas Libraries' Experience.”**

Gelaw, Daniel, S.K. Hastings, and Cathy Hartman. *First Monday* 7, no. 8 (August 2002).

<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/981>.

This paper discusses the issues related to digital resources preservation and demonstrates the role of preservation metadata in facilitating the preservation activities in general.

**“Comparison of Strategies and Policies for Building Distributed Digital Preservation Infrastructure: Initial Findings from the MetaArchive Cooperative.”**

Halbert, Martin. *International Journal of Digital Curation* 4, no. 2 (October 15, 2009).

<http://www.ijdc.net/index.php/ijdc/article/view/117>.

This paper discusses distributed, cooperatively maintained approaches to digital content preservation infrastructures for cultural memory organizations (CMOs). Initial findings are presented from an early organizational effort (the MetaArchive Cooperative).

**“Management and preservation of research data with iRODS.”**

Hedges, Mark, Adil Hasan, and Tobias Blanke. In *Proceedings of the ACM first workshop on CyberInfrastructure: information management in eScience*, 17-22. Lisbon, Portugal: ACM, 2007.

<http://portal.acm.org/citation.cfm?id=1317353.1317358>.

This paper presents first steps towards implementing a data layer to support a semi-automated preservation management system for research data in the arts and humanities.

**“Rule-based curation and preservation of data: A data grid approach using iRODS.”**

Hedges, Mark, Tobias Blanke, and Adil Hasan. *Future Generation Computer Systems* 25, no. 4

(April 2009): 446-452. <http://www.sciencedirect.com/science/article/B6V06-4TPF47W->

[3/2/df291f8d67288a87bf7355b251a2a0ef](https://doi.org/10.2196/3/2/df291f8d67288a87bf7355b251a2a0ef).

Describes an approach to automation of curation, in which digital curation policies and strategies are represented as rules, which are implemented in data grids based on the iRODS middleware.

**“The DCC Curation Lifecycle Model.”**

Higgins, Sarah. *International Journal of Digital Curation* 3, no. 1 (December 2, 2008).

<http://www.ijdc.net/index.php/ijdc/article/view/69>.

Outlines the DCC Curation Lifecycle Model and the uses for which it is intended.

**“Relay-supporting Archives: Requirements and Progress.”**

Janeé, Greg, James Frew, and Terry Moore. *International Journal of Digital Curation* 4, no. 1 (June 29, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/102>.

This article characterizes long-term preservation of digital content as an extended relay in time, in which repeated handoffs of information occur: it examines the support of current preservation technologies for these handoffs, notes shortcomings, and suggests some improvements.

**“Sun Preservation and Archive Special Interest Group: May 2008 Meeting.”**

Mays, Vicky, and Ian Dolphin. *Ariadne* 56 (July 2008). <http://www.ariadne.ac.uk/issue56/pasig-2008-05-rpt/>.

Review of the Sun Preservation and Archiving Special Interest Group (PASIG) meeting in San Francisco of 27-29 May 2008.

**“Towards a Theory of Digital Preservation.”**

Moore, Reagan. *International Journal of Digital Curation* 3, no. 1 (December 2, 2008).

<http://www.ijdc.net/index.php/ijdc/article/view/63>.

Examines the representation information that is needed about the preservation environment for a theory of digital preservation, including descriptions of the preservation management policies, the preservation processes, and the state information that is needed to verify the correct working behavior of the system.

**“DCC Workshop Report: E-mail Curation: Practical Approaches for Long-term Preservation and Access, Newcastle-upon-Tyne, April 24 - 25, 2006.”**

Pennock, Maureen. *International Journal of Digital Curation* 1 (2006).

<http://www.ijdc.net/index.php/ijdc/article/view/11>.

A report on the Digital Curation Centre workshop held in Newcastle-upon-Tyne in April 2006 to explore practical approaches for managing, preserving and re-using e-mail records.

**“Data management and curation.”**

Research Information Network. <http://www.rin.ac.uk/our-work/data-management-and-curation>.

Resource relating to 'arrangements for the creating, sharing and use of research data and related policy implications' from the RIN website.

**“The Data Curation Continuum: Managing Data Objects in Institutional Repositories.”**

Treloar, Andrew, David Groenewegen, and Cathrine Harboe-Ree. *D-Lib Magazine* 13, no. 9 (9,

2007). <http://www.dlib.org/dlib/september07/treloar/09treloar.html>.

This article introduces the idea of a Data Curation Continuum and describes the various continua that might be applicable in a repository data management context, and discusses some of the implications of this approach.

#### **“Digital curation toolbox.”**

Wellcome Library. <http://library.wellcome.ac.uk/node289.html>.

This page contains some of the documents that have been produced by the Wellcome Library to explain how we work with born digital material and the processes of digital curation.

## **Repositories**

#### **“SWORD: Simple Web-service Offering Repository Deposit.”**

Allinson, Julie, Sebastien Francois, and Stuart Lewis. *Ariadne* 54 (January 2008).

<http://www.ariadne.ac.uk/issue54/allinson-et-al/>.

Describes the JISC-funded SWORD Project which has produced a lightweight protocol for repository deposit.

#### **“Versioning in Repositories: Implementing Best Practice.”**

Brace, Jenny. *Ariadne* 56 (July 2008). <http://www.ariadne.ac.uk/issue56/brace/>.

Explains why giving time to versioning within a repository is worthwhile and outlines the best practice to implement.

#### **“RepoMMan: Delivering Private Repository Space for Day-to-day Use.”**

Green, Richard, and Chris Awre. *Ariadne* 54 (January 2008).

<http://www.ariadne.ac.uk/issue54/green-awre/>.

Describe work undertaken at the University of Hull to place Web services at the heart of its personalised digital repository provision.

#### **“The REMAP Project: Steps towards a Repository-enabled Information Environment.”**

Green, Richard, and Chris Awre. *Ariadne* 59 (April 2009).

<http://www.ariadne.ac.uk/issue59/green-awre/>.

Investigates what role a repository can play in enabling and supporting the management and preservation of its own digital content.

#### **“Digital Preservation Service Provider Models for Institutional Repositories.”**

Hitchcock, Steve, Tim Brody, Jessie M.N. Hey, and Leslie Carr. *D-Lib Magazine* 13, no. 5 (5, 2007). <http://www.dlib.org/dlib/may07/hitchcock/05hitchcock.html>.

This article describes the evolution of a series of models that have informed progress towards the conception of flexible and distributed preservation services for institutional repositories.

#### **“Digital preservation in the context of institutional repositories.”**

Hockx-Yu, Helen. *Program: electronic library and information systems* 40, no. 3 (2006): 232-243.

<http://www.emeraldinsight.com/journals.htm?articleid=1563488&show=pdf>.

This paper will help the reader to gain a better understanding of the issues related to digital preservation in general and how JISC's work has helped to tackle these issues. This paper states JISC's view on, and future plan for, digital repositories.

**“Motivating and Impeding Factors Affecting Faculty Contribution to Institutional Repositories.”**

Kim, Jihyun. *Journal of Digital Information* 8, no. 2 (September 26, 2007).

<http://journals.tdl.org/jodi/article/viewArticle/193>.

Faculty contribution is considered one of the success factors for an IR even though several studies have found low rates of faculty submission. In order to learn how we might be able to address this problem, this study investigated factors that motivate or impede faculty contribution, via a survey of a sample of 67 professors.

**“Just email it to me!: why things get lost in shared file repositories.”**

Rader, Emilee. In *GROUP '07 Doctoral Consortium papers*, 1-2. Sanibel Island, Florida: ACM, 2007.

<http://portal.acm.org/citation.cfm?id=1329121&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

The use of shared file repositories in organizations is becoming more frequent; however, repository users are not always able to effectively find and access information. This article documents and analyzes the scope and consequences of the problem, and tests hypotheses about possible remedies.

**“Repositories Support Project Publications and Podcasts,”**

Repositories Support Project. <http://www.rsp.ac.uk/pubs/>.

This page gives details and links for publications, presentations and podcasts prepared by RSP partners and/or written about the Repositories Support Project.

**“If you build it, will it fly? Criteria for success in a digital repository.”**

Thibodeau, Kenneth. *Journal of Digital Information* 8, no. 2 (September 26, 2007).

<http://journals.tdl.org/jodi/article/viewArticle/197>.

International collaborations have produced a standard describing the functions of a digital repository and the characteristics of one that can be trusted. These results provide an abstract frame of reference for evaluating such repositories, but meaningful evaluation requires that they be supplemented by empirical data on the purpose of each repository and the institutional, cultural and resource context in which it operates.

## Personal Information Management

**“Finding and reminding: file organization from the desktop.”**

Barreau, Deborah, and Bonnie A. Nardi. *SIGCHI Bull.* 27, no. 3 (1995): 39-43.

<http://portal.acm.org/citation.cfm?id=221307&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

This paper summarizes and synthesizes two independent studies of the ways users organize and find files on their computers.

**“Plenty of Room at the Bottom? Personal Digital Libraries and Collections.”**

Beagrie, Neil. *D-Lib Magazine* 11, no. 06 (6, 2005).

<http://www.dlib.org/dlib/june05/beagrie/06beagrie.html>.

People are capturing and storing an ever-increasing amount of digital information about or for themselves, and many issues arise from this. This article provides a commentary on current

research and emerging services in this area and discusses potential implications for individuals, libraries and their institutions.

**“Information scraps: How and why information eludes our personal information management tools.”**

Bernstein, Michael, Max Van Kleek, David Karger, and M. C. Schraefel. *ACM Trans. Inf. Syst.* 26, no. 4 (2008): 1-46.

<http://portal.acm.org/citation.cfm?id=1402263&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

Investigates information scraps: content scribbled on Post-it notes, sent in email messages to ourselves, and so on. Looks at why we have them, the role they play, and how PIM applications might be better designed to accommodate and support them.

**“Personal Information Management Bibliography,”**

Boardman, Richard. 2003. <http://www.iis.ee.ic.ac.uk/~rick/pim-biog.htm>.

Covers studies of PIM, theories/methods/background, and tool design/evaluation.

**“‘Stuff goes into the computer and doesn't come out’: a cross-tool study of personal information management.”**

Boardman, Richard, and M. Angela Sasse. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 583-590. Vienna, Austria: ACM, 2004.

<http://portal.acm.org/citation.cfm?id=985766&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

This paper reports a study of Personal Information Management (PIM), which advances research in two ways: (1) rather than focusing on one tool, we collected cross-tool data relating to file, email and web bookmark usage for each participant, and (2) we collected longitudinal data for a subset of the participants.

**“Information behaviour that keeps found things found.”**

Bruce, Harry, William Jones, and Susan Dumais. <http://informationr.net/ir/10-1/paper207.html>.

This paper reports on a study that the researchers call: 'Keeping found things found on the Web' or 'KFTF'. The research focuses on the classic problem of ensuring that once a useful information source or channel has been located, it can be found again when it is needed.

**“On Understanding the Relationship Between Recollection and Refinding.”**

Elsweiler, David, Mark Baillie, and Ian Ruthven. *Journal of Digital Information* 10, no. 5 (November 5, 2009). <http://journals.tdl.org/jodi/article/viewArticle/436>.

Focusing on email re-finding, this article investigates the relationship between recollection and information re-finding performance. We discuss our findings with respect to past and future work and also to the design of new re-finding tools.

**“Finding and reminding' reconsidered.”**

Fertig, Scott, Eric Freeman, and David Gelernter. *SIGCHI Bull.* 28, no. 1 (1996): 66-69.

<http://portal.acm.org/citation.cfm?id=249187&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

### **“Students' Organization Strategies of Personal Information Space.”**

Hardof-Jaffe, Sharon, Arnon HersHKovitz, Hama Abu-Kishk, Ofer Bergman, and Rafi Nachmias. *Journal of Digital Information* 10, no. 5 (November 5, 2009).

<http://journals.tdl.org/jodi/article/viewArticle/438>.

This study uses new powerful data mining methods in order to observe students' personal information space organization strategies in the personal Web space allocated to them on the university servers.

### **“How do people organize their desktops?”**

Henderson, Sarah. In *CHI '04 extended abstracts on Human factors in computing systems*, 1047-1048. Vienna, Austria: ACM, 2004.

<http://portal.acm.org/citation.cfm?id=985972&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

This research aims to understand how people currently organize their documents, identify the strengths and weaknesses of current systems and explore the usefulness of other information structures.

### **“Being Wired or Being Tired: 10 Ways to Cope with Information Overload.”**

Houghton-Jan, Sarah. *Ariadne* 56 (July 2008). <http://www.ariadne.ac.uk/issue56/houghton-jan/>.

Explores different strategies for managing and coping with various types of informational overload.

### **“A Review of Organizational Structures of Personal Information Management.”**

Indratmo, J., and Julita Vassileva. *Journal of Digital Information* 9, no. 1 (May 19, 2008).

<http://journals.tdl.org/jodi/article/viewArticle/251>.

Synthesizes and classifies existing research on PIM based on the approach used to organize information items. The article discusses the strengths and weaknesses of different approaches along with examples showing how to deal with the weaknesses, and provides design recommendations and a framework for researchers to experiment with various ideas for developing novel PIM tools.

### **“Digital management and retrieval of physical documents.”**

Jervis, Matthew, and Masood Masoodian. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, 47-54. Cambridge, United Kingdom: ACM, 2009.

<http://portal.acm.org/citation.cfm?id=1517682&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

This paper presents an architecture for a system to digitally manage paper document containers, such as folders, which has been used as the design basis for a prototype system we have developed.

### **“Finders, keepers? The present and future perfect in support of personal information management.”**

Jones, William. *First Monday* 9, no. 4 (March 2004).

<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1123>.

Discusses making decisions about how much and what information to keep, and tools that might help people do this.

**“Personal Information Management.”**

Jones, William. *Annual Review of Information Science and Technology* 41, no. 1 (2007): 453-504.  
[http://www3.interscience.wiley.com/cgi-bin/fulltext/121477218/main.html,ftx\\_abs](http://www3.interscience.wiley.com/cgi-bin/fulltext/121477218/main.html,ftx_abs).

Covers: influences on PIM; analysis of PIM; research to understand how people do PIM; methodologies of PIM inquiry; approaches to PIM integration; information fragmentation.

**“Don’t Take My Folders Away! Organizing Personal Information to Get Things Done.”**

Jones, William, Ammy Jiranida Phuwannurak, Rajdeep Gill, and Harry Bruce.

<https://digital.lib.washington.edu/dspace/handle/1773/2031>.

A study exploring the way people organize information in support of projects (“teach a course”, “plan a wedding”, etc.), focusing particularly on the use of folders.

**“Using memory for events in the design of personal filing systems.”**

Lansdale, Mark, and Ernest Edmonds. *Int. J. Man-Mach. Stud.* 36, no. 1 (1992): 97-126.

<http://portal.acm.org/citation.cfm?id=141734&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

This paper describes a prototype interface, MEMOIRS (Managing Episodic Memory for Office Information Retrieval Systems), which is designed to support the management of personal information in a new way. This approach treats a personal filing system as a history of events (of which documents are a particular type), and focuses upon user's recall for those events.

**“Personal digital library: pim through a 5s perspective.”**

Ma, Yi, Edward A. Fox, and Marcos A. Gonçalves. In *Proceedings of the ACM first Ph.D. workshop in CIKM*, 117-124. Lisbon, Portugal: ACM, 2007.

<http://portal.acm.org/citation.cfm?id=1316893&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

In this paper, we present a formal framework for personal information management through a personal digital library perspective based upon the 5S approach. This framework includes a formal definition of the components and functionalities for a minimal personal digital library. We believe this is one of the first efforts to provide a theoretical foundation for personal information management.

**“How do people organize their desks?: Implications for the design of office information systems.”**

Malone, Thomas W. *ACM Trans. Inf. Syst.* 1, no. 1 (1983): 99-112.

<http://portal.acm.org/citation.cfm?id=357430&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

**“A 'pile' metaphor for supporting casual organization of information.”**

Mander, Richard, Gitta Salomon, and Yin Yin Wong. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 627-634. Monterey, California, United States: ACM, 1992.

<http://portal.acm.org/citation.cfm?id=143055&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

Describes the development and testing of a new desktop interface element, the pile, arising from a user study investigating how people deal with the flow of information in their workspaces.

**“Beyond paper: supporting active reading with free form digital ink annotations.”**

Schilit, Bill N., Gene Golovchinsky, and Morgan N. Price. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 249-256. Los Angeles, California, United States: ACM Press/Addison-Wesley Publishing Co., 1998.

<http://portal.acm.org/citation.cfm?id=274680&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

**“PaperSpace: a system for managing digital and paper documents.”**

Smith, Jeff, Jeremy Long, Tanya Lung, Mohd M. Anwar, and Sriram Subramanian. In *CHI '06 extended abstracts on Human factors in computing systems*, 1343-1348. Montréal, Québec, Canada: ACM, 2006.

<http://portal.acm.org/citation.cfm?id=1125700&dl=GUIDE&coll=GUIDE&CFID=97795080&CFTOKEN=56870801>.

Here we present PaperSpace, a computer vision based document management system that allows users to combine paper and digital documents. Using PaperSpace users can locate paper copies of printed digital documents, retrieve digital versions of paper documents and fluidly move between digital and paper documents.

**“The character, value, and management of personal paper archives.”**

Whittaker, Steve, and Julia Hirschberg. *ACM Trans. Comput.-Hum. Interact.* 8, no. 2 (2001): 150-170. <http://portal.acm.org/citation.cfm?doid=376929.376932>.

We explored general issues concerning personal information management by investigating the characteristics of office workers' paper-based information, in an industrial research environment, and investigating the merits of different paper-processing strategies such as filing and piling.

**“Email overload: exploring personal information management of email.”**

Whittaker, Steve, and Candace Sidner. In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, 276-283. Vancouver, British Columbia, Canada: ACM, 1996.

[http://portal.acm.org/ft\\_gateway.cfm?id=238530&type=html&coll=GUIDE&dl=GUIDE&CFID=97795080&CFTOKEN=56870801](http://portal.acm.org/ft_gateway.cfm?id=238530&type=html&coll=GUIDE&dl=GUIDE&CFID=97795080&CFTOKEN=56870801).

**“The personal curation of digital objects: A lifecycle approach.”**

Williams, Peter, Jeremy Leighton John, and Ian Rowland. *Aslib Proceedings* 61, no. 4 (2009): 340-363. <http://www.emeraldinsight.com/journals.htm?articleid=1800798&show=abstract>.

This paper aims to set out a coherent intellectual framework to help to better understand how people create, organise, manage, use and dispose of their personal digital archives.

## Metadata and Related Issues

**“Nine questions to guide you in choosing a metadata schema.”**

Kennedy, Marie D. *Journal of Digital Information* 9, no. 1 (May 19, 2008).

<http://journals.tdl.org/jodi/article/viewArticle/226>.

This article is a guide for collection developers at the point of considering a metadata schema

for their digital collection. The nine questions asked in this article will assist a developer in clarifying how he wants the collection to be organized, described, and used.

**“A Metadata Registry for Metadata Interoperability.”**

Li, Jian-hui, Jia-xin Gao, Ji-nong Dong, Wei Wu, and Yan-fei Hou. *Data Science Journal* 6 (2007). [http://www.jstage.jst.go.jp/article/dsj/6/0/6\\_S379/article](http://www.jstage.jst.go.jp/article/dsj/6/0/6_S379/article).

A metadata registry is a place to keep facts about characteristics of data that are necessary for data sharing and exchange in a specific domain. This paper will explore the role of metadata registries and describe some of the experiences of implementing the registry.

**“Policy-based Distributed Data Management Systems.”**

Rajasekar, Arcot, Reagan Moore, Mike Wan, and Wayne Schroeder. *Journal of Digital Information* 11, no. 1 (March 22, 2010). <http://journals.tdl.org/jodi/article/viewArticle/756>.

Scientific research collaborations generate massive amounts of data. Policy-based data management systems minimize the amount of labor needed to manage the massive collections by automating the enforcement of management policies and the validation of assessment criteria.

**“Embedding Metadata and Other Semantics in Word Processing Documents.”**

Sefton, Peter, Ian Barnes, Ron Ward, and Jim Downing. *International Journal of Digital Curation* 4, no. 2 (October 15, 2009). <http://www.ijdc.net/index.php/ijdc/article/view/121>.

This paper describes a technique for embedding document metadata, and potentially other semantic references inline in word processing documents, which the authors have implemented with the help of a software development team.

**“Development of a Science Database System Applicable to Various Access Restrictions.”**

Takata, Y., Y. Kasahara, and T. Matsuhira. *Data Science Journal* 8 (2009). [http://www.jstage.jst.go.jp/article/dsj/8/0/8\\_IGY32/article](http://www.jstage.jst.go.jp/article/dsj/8/0/8_IGY32/article).

Describes the development of a general-purpose Web-based database (DB) system that manages and opens experimental and/or observational data accumulated in universities and academic institutes.

**“Spinning a Semantic Web for Metadata: Developments in the IEMSR.”**

Tonkin, Emma, and Alexey Strelnikov. *Ariadne* 59 (April 2009). <http://www.ariadne.ac.uk/issue59/tonkin-strelnikov/>.

Reflects on the experience of developing components for the Information Environment Metadata Schema Registry.

**“Dataset acquisition, accessibility, annotation, e-research technologies (DART) project.”**

Tsoi, Ah Chung, Jeff McDonell, Andrew Treloar, and Ian Atkinson. *International Journal on Digital Libraries* 7, no. 1 (October 1, 2007): 53-55. <http://dx.doi.org/10.1007/s00799-007-0019-4>.

The DART project undertook a coordinated program of e-Research requirements analysis, software development, policy and guideline creation and prototyping to investigate how best to deal with various issues raised by large data sets and streams.

**“Imagery metadata development based on ISO/TC 211 standards.”**

Xie, Rong, and Ryosuke Shibasaki. *Data Science Journal* 6 (2007): 28-45.

[http://www.jstage.jst.go.jp/article/dsj/6/0/6\\_28/article](http://www.jstage.jst.go.jp/article/dsj/6/0/6_28/article).

This paper reviews the present status and major problems of the existing ISO standards related to imagery metadata. An imagery metadata model is proposed to facilitate the development of imagery metadata on the basis of conformance to these standards and combination with other ISO standards related to imagery.