

## **INF 389G Introduction to Electronic and Digital Records, Unique # 28900, Fall 2014**

Professor: Dr. Patricia K. Galloway

Course Meeting Times: Wednesday 12:00 p.m - 3:00 p.m, UTA 1.201A

### **Course Description**

The management, preservation, and use of electronic records and other digital objects with enduring (or even temporary) value are almost all still problems with only partial solutions. Although increasing progress is being made by archival researchers and some standards have emerged for “simple” digital records, there are two reasons why this open-ended situation will probably remain constant: the supporting technologies are changing constantly and the rate of change is accelerating; and creators and users of these records (if not the records' potential managers and preservers) are themselves caught up in a culture of immediacy that makes the problems with electronic records invisible until some legal entanglement brings them into sharp focus (as, for example, the destruction of records by Enron, 9/11 terrorists, and the collection of emails by the NSA)—or you suddenly realize that you have lost the only digital photos you still had of some crucial event in your life.

Yet since both governments and other human institutions and individuals have depended upon technologies of memory in the past, it is a safe bet that they will continue to do so in at least the immediate future (as, for example, Barack Obama's Blackberry). For that reason these problems must and will be solved, at least in terms of a sequence of temporary solutions that will be good enough to achieve the ends of the institutions in question and of individuals for their everyday lives, both by those who are charged with the institutional custody and preservation of the cultural record and by individuals themselves.

The problems are not just technological; if that were so they could (and perhaps would) already have been solved. They are, more importantly, social, economic, and political. The archivist or records manager or digital librarian or individual called upon to solve them in a real-world setting will have to understand not just a set of ideal archival requirements, but how to cope with applying them to and tailoring them for an actual functional environment, one where change never ceases, where the people who create and use the records have other things to think about, where the powers that be continue to think of the problem as the job of IT, and where getting it right once and for all is not an option. Individuals can hopefully borrow from these institutional practices the solutions that suit them—or they may devise novel solutions for themselves. Increasingly, it seems that individual practices are having significant impact on what people can be persuaded to do in the way of digital recordkeeping in the workplace (especially where BYOD is becoming common), so personal digital archives is becoming an important area of research.

In this introductory course, we will become acquainted with the basic literature on digital records and recordkeeping (and contest the term “records”), track new developments in the field in order to get a feel for how to pay attention to emergent problems, apply our learning to coming up with a business case for a real-world problem, and examine reflexively our own digital recordkeeping



practices over our lifetimes and at present as a sample of the kinds of problems existing in the broader environment.

Professor: Dr. Patricia K. Galloway  
Email: galloway@ischool.utexas.edu  
Phone: (512) 232-9220  
Office: UTA 5.436  
Office Hours: by appointment

Teaching Assistant: Pearl Ko  
Email: pearlyko@utexas.edu  
Contact Information: Please contact by email or by appointment

## Course Objectives

In this course we will address in readings digital records and digital records practices in general in order to flush out legal requirements for authenticity, but we will use our own personal digital records as a leverage tool to help understand the relevant principles and varieties of digital recordkeeping problems that people encounter every day at home and at work. We will be concerned with the issues that cluster around the creation, acquisition, preservation, and use of digital records. To be more specific, we will attempt to:

- Understand how digital records exist and function as records (or nonrecords: who gets to decide? and what are digital records anyway?)
- Why people create them and use them as they do in their original functional environment; why people need and keep them; why people actively destroy them or allow them to die
- Understand the implications of both statutory requirements and the technological environment for digital recordkeeping
- Review major trends in digital records archivy
- Appreciate and evaluate approaches to problems of media obsolescence
- Learn how to monitor and anticipate the digital recordkeeping implications of technological change; be familiar with both traditional and emerging digital genres
- Participate in the identification, acquisition, and management of digital records

There is a textbook this year that you should be able to order on the day of the first class and have in hand by the time you need it, but most other readings will be available online; if you should encounter difficulty accessing them, please contact our TA immediately, and NOT the day before class. The syllabus will be posted and should be checked regularly for any changes. Any students requiring accommodation for disability or religious holy days should contact me at the beginning of class.

## Textbook

The assigned textbook is *Personal Information Management*, edited by William Jones and Jaime Teevan (University of Washington Press, 2007), available at the COOP or through the usual online vendors in new or used condition from around \$12-\$30-something.



The essay below is a little outdated, but the archival perspective is sound and it is still worth reading:

Anne J. Gililand-Swetland, *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment* (Washington: CLIR, 2000). Available online at <http://www.clir.org/pubs/reports/pub89/contents.html> You can purchase it online from the CLIR site if you want a nice paper copy.

Other readings are available online (on the Internet or through the Library) or through Canvas.

If you feel you need more background in any specific area (archives, technology, etc.), check the Resources page; and if you find a valuable reading, please share it with the class and we'll add it to the page.

## **Assignments**

This class will generally be run like a seminar, with discussions focusing on the assigned readings, assigned student presentations on recent technology items, and tasks carried out outside of class. There will be lectures and discussion in class on key areas of interest in digital records throughout the course. Course requirements consist of class attendance and participation, including active contributions to class discussions, a recent technology presentation, the preparation of a personal digital records management plan, and the preparation of a digital records business plan as a member of a small team.

### **Class participation (15% of grade)**

Class attendance is expected for all students barring illness or religious holy days (please contact me regarding these); multiple absences will affect a student's final grade for the course. Each student is expected to complete the required readings on a weekly basis and be able to discuss them in class, to demonstrate that the reading has been done and understood. Note that specific points for discussion that you are expected to prepare have been provided, though you are welcome to raise additional issues. This is a serious requirement and can make a letter grade's difference (do the math if you are concerned about grades). We have fifteen class meetings; the first class meeting will be taken up with a preliminary discussion of the topic of the course, while the last will include presentations of the team business plans, leaving twelve classes in which we will discuss readings, current developments in the information realm that affect the creation and management of digital objects, and, once the semester has begun, progress on both your digital preservation business case and your personal digital archive management plan. The fifteenth scheduled class time on November 12 will be used for some digital visualization exercises.

### **Recent technology report (15% of grade)**

For eight of our discussion classes, we will have a presentation on a new technology that has emerged relatively recently. Small student groups will be assigned a topic relating to a new technology to research (find out if it is legitimate, find out how successful it is, test it out if appropriate, then seriously figure out what kind of digital objects are created by or otherwise



concerned in its use, what happens to them, who keeps them, etc. etc.) and **all students assigned to each technology report will present in class**. You will talk about the subject for ten minutes (ppt presentations are not required, but you do need to present the technology in some way--talk, call up pages from the Internet, and/or stage demonstrations if appropriate) and then you will lead the class in discussing the impact of the technologies on recordkeeping--their own and that in other settings (so you'll need to prepare some questions to spark discussion). Each student team will provide a "documentary package" to be added to the Resources page for other students to have access to (for example, a copy of the ppt presentation you gave, a set of URLs linking to websites relating to the topic, relevant readings you have found, etc.).

### **Digital Preservation Business Case Planning Project (35% of grade)**

Each student will be part of a team working on a planning project for digital preservation to solve a real-world problem for a real organization. There are four projects and the teams will consist of 4-5 people. We will use a planning instrument ([http://wiki.dpconline.org/index.php?title=Digital\\_Preservation\\_Business\\_Case\\_Toolkit](http://wiki.dpconline.org/index.php?title=Digital_Preservation_Business_Case_Toolkit)) developed by the SPRUCE project in the UK, drawing on the business case literature and their intimate knowledge of the problems of digital preservation. The project will be divided up into sections within the teams and students will report out both separately and as a team. The four projects for this semester are:

- A plan for preserving the existing historical syllabi for past iSchool classes currently linked to the iSchool website
- A plan for processing an accession of born-digital video from UT University Communications for the Briscoe Center
- A plan for inventorying and archiving the born-digital records of the CHIPS student group
- TBA

### **Digital Preservation Business Cases are due on November 26.**

### **Personal Digital Archive Management Plan (35% of grade):**

Each student will write a report outlining a personal digital archive management plan developed during the semester. The framework for this plan will be discussed in class and a handout will be provided early on. We will also discuss your progress and we will workshop some of the elements of the plan in class as the semester progresses.

### **Personal Digital Archive Plans are due on December 3.**



## Schedule

**NOTE: This schedule is definitely preliminary until after the first class meets and may change slightly through the semester if new issues come up--so don't just print or download it and continue to refer to that version.** URLs change constantly, so if you find a dead link please do **both** of the following two things:

- 1) Stick the URL into the Wayback Machine at <http://www.archive.org> and see if you can find it there; and
- 2) Let our TA know one way or the other: if you found it, the Wayback URL for it; or if you didn't find it, so we can do something about it.

### August 27: Background, overview of the field

Reminder about ordering **textbook**: Jones and Teevan, *Personal Information Management*.

Fill out **questionnaire** about educational, technical, and archival back ground (in class)

Discuss **course and course assignments**, including:

- New Technology research assignment (15%)
- Personal digital archive management report (35%) Note that if you begin thinking about this project from the beginning, you can make use of class discussions to explore and sharpen your thinking about your own records
- Digital Preservation Business Case project (35%)
- Class participation (15%)

**Topic:** Overview of the field of digital recordkeeping—history and areas of interest in a changing social and communication environment. Lecture-discussion. Be prepared to participate by sharing some of your own experiences with digital materials.

### September 3: What is a digital record and how can I deal with it?

**New Technology** research assignments handed out and discussed.

**Topic:** Definitions of "electronic and digital records" and the range of objects and environments that are implicated under this rubric. Discuss the archival view of digital records and the skills that seem to be required for coping with them. Students will be expected to have read the assignments and to be prepared to discuss them critically. For a start, read the readings and then prepare to discuss the questions below—which means: have an answer in mind and/or written down with your preparation notes for class.

#### Questions to prepare for discussion:

- 1) What do you consider the most valuable part of the archival perspective as outlined by Gilliland? Do you consider that any parts are now outdated, and if so why?



2) How does your skillset fit with the New Skills inventory in general and as expanded in detail in the DigCCurr matrix? If you have questions about the matrix and the skills, please raise them. How might you update or add to your skills?

### ***Readings***

Anne J. Gilliland-Swetland, *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment* (Washington: CLIR, 2000)

<http://www.clir.org/pubs/reports/pub89/pub89.pdf>

Richard Pearce-Moses and Susan E. Davis, "Knowledge and Skills Inventory," in *New Skills for a Digital Era* (2006), pp. 1-31, available at

<http://www.archivists.org/publications/proceedings/NewSkillsForADigitalEra.pdf> The proceedings of this conference are worth reading as a whole if and when you have time, not least because three of the case studies are from the UT iSchool.

Cal Lee, *Matrix of digital curation knowledge and competencies*, 2009. Available at <http://www.ils.unc.edu/digccurr/digccurr-matrix.html> BE SURE TO DRILL DOWN INTO THE DIMENSIONS!!

Recent job descriptions for digital preservation workers:

<http://www.dpconline.org/newsroom/vacancies/1247-vacancy-at-archives-new-zealand-digital-preservation-process-administrator>

### **September 10 : Official digital records and regulation by statute and computer code**

Protocol for preparation of **Personal Information Management Plan** discussed.

**Topic:** How people use and work with electronic records/digital objects, including differences that may be introduced by the electronic medium; the electronic environment and how it is literally legislated from scratch by computer code (and the implications of "net neutrality"), how real legislation deals with paper and digital records, and how individuals manage their own records. Review the MDAH digital archives project and how it evolved 1999-2010 (including Sovereignty Commission migration and Walker Sampson's work on moving digital archives to DSpace: <http://wsampson.wordpress.com/> October 12); also discuss the TERM email project and its "failure" (see October 8 also for email), as well as the new 2014 TSLAC digital archives initiative. How "record" and "evidence" are constituted and how hard it is to do this.

### **Questions to prepare for discussion:**

1) What digital records should governments keep? Think of this in terms of a specific issue that interests you (Social Security, taxes, passports, gun licenses, etc.) and find out what the government of Texas actually requires for our discussion.



2) How do the realities of computer technology make it possible (or not) to **keep** digital records? You might work through this by asking what problems NARA has had in developing a digital archive (see links in the last reading). If you are interested in more detail about NARA's long struggle with e-records, see Bruce Ambacher (ed.), *Thirty Years of Electronic Records*.

### ***Readings***

Lawrence Lessig, *Code and other Laws of Cyberspace*, Part 1 (1999). Available on Canvas in two files, 1.1 and 1.2 , under Course Documents. This is version 1, which Lessig refers to now as an "ancient text"; the revised version, revised via wiki, can be found as a free download or wiki at <http://codev2.cc/> (you can add your own remix). If Lessig's work on how the digital environment changes the impact and meaning of traditional legal frameworks interests you, succeeding books include *The Future of Ideas*, *Free Culture*, and *Remix*.

Uniform Electronic Transactions Act (UETA) summary: this act, passed in 1999, first made digital records officially acceptable in legal transactions. Read a short general summary at: <http://www.ncsl.org/default.aspx?tabid=13484>

"Electronic Records standards and procedures," from TSLAC, revised 2005:

<http://www.tsl.state.tx.us/slr/recordspubs/stbull01.html>

Also read the Texas public records statutes at Government Code, chapters 441.180-197 (Texas State Library and management of records), 551.021-023 (Open meetings records), 552 (Public Information law: especially subsections 101-136 listing exceptions and 272 on access to electronic records): all these can be found at <http://www.statutes.legis.state.tx.us>

For a list of all the national and international standards that pertain to the official handling of digital records, see a listing here (removed from the DIR site, now only on the Internet Archive): <http://web.archive.org/web/20100616080137/http://www.dir.state.tx.us/pubs/derm/standards/section1.htm>

Mississippi Department of Archives and History Electronic Records Project, [Management Standards](#) (1999).

State Electronic Records Initiative (SERI; sponsored by CoSA):

<http://www.statearchivists.org/seri/> This project began in 2011 and is designed to run through 2015 with different aspects of the project.

National Research Council, Building an Electronic Records Archive at the national Archives and Records Administration: Recommendations for a Long-Term Strategy (2005). This was the attempt to bring the project up to speed after long delay. It's available freely downloadable from here: [http://www.nap.edu/openbook.php?record\\_id=11332&page=R1](http://www.nap.edu/openbook.php?record_id=11332&page=R1)

Federal News Radio, "NARA to suspend development of ERA starting in 2012," 12/17/2010, located here: <http://www.federalnewsradio.com/?sid=2204570&nid=697>

**September 17: Personal digital information and regulation by computer code**



## **New Technology presentation 1:**

**Digital Preservation Business Case Planning Project** assignments will be handed out.

**Topic:** Non-official born-digital (and born-again-digital) objects and how they are managed and preserved. Discuss digital personal records, with a focus on the student project (assignment handed out September 3) to understand one's own digital belongings and manage a personal archive. Further discussion of the instrument students are using to formulate their personal digital archives management plan.

**Questions to prepare for discussion (use examples both from the readings for today and from your thinking about your own records):**

- 1) How do the digital tools a person personally uses constrain what they create and what they can personally keep?
- 2) Is it possible and/or desirable to know and manage the full range of digital objects that a person presently creates? Think about an example or two from the readings and your own experience and why it might be complicated to do so.
- 3) Is it possible and/or desirable to find out and understand how one's identity is represented on the Internet? How is this a new problem?
- 4) Should archivists or commercial vendors assist ordinary people to manage their digital belongings? Think of arguments in favor of each alternative.

### ***Readings***

*Personal Information Management*, ed. William Jones and Jaime Teevan (hereafter cited as *PIM*), pp. 3-75, chapters 1-4 will begin our investigation of how people actually keep digital records, regardless of statutes.

Neil Beagrie, "Plenty of Room at the Bottom? Personal Digital Libraries and Collections," *D-Lib* 11(6), June 2005, available at <http://www.dlib.org/dlib/june05/beagrie/06beagrie.html>

Paradigm Project Workbook on Digital Private Papers (2006; read Introduction and Collection Development sections): <http://www.paradigm.ac.uk/workbook/index.html> This reading will provide you with some additional tools for your personal digital archiving project.

Peter Williams, Jeremy Leighton John, and Jan Rowland, "The Personal Curation of Digital Objects: a Lifecycle Approach," *ASLIB Proceedings* 61(4), 2009, 340-363. Available under "journals" from the PCL catalog.

Simson Garfinkel and David Cox, "Finding and Archiving the Internet Footprint" (2009): <http://simson.net/clips/academic/2009.BL.InternetFootprint.pdf>



Clive Thompson, "A Head for Detail," Fast Company 110 (November 2006): <http://www.fastcompany.com/magazine/110/head-for-detail.html> This essay is an entertaining account of Gordon Bell's "life-logging" experiment, which signaled Microsoft's interest in being involved in this space.

## **September 24: Record granularity and metadata**

### **New Technology presentation 2:**

**Topic:** Implications for records creators, archives, and users of record-level description and the generation of metadata to provide it. Discuss the issue of descriptive granularity and review various metadata schemes. Investigate in class metadata created by programs.

### **Questions to prepare for discussion:**

- 1) How does the need for bitstream-level metadata contradict or make problematic the adoption of minimal processing (aka MPLP) standards in an archive? Is there a solution to this contradiction?
- 2) What kinds of metadata are needed for keeping your own records? How does this differ (if it does) from the kinds needed for archival collections? For digital library collections?
- 3) Give/find an example of metadata among your own digital files.

### ***Readings***

David Bearman, "Item Level Control and Electronic Recordkeeping" (this is a classic 1996 article that makes a very important point while summarizing the Pittsburgh project; the entire Pittsburgh project website, by the way, had not been archived and was lost by a site remodel but has been recovered and restored by students at the Pittsburgh School of Information; there is a link to it on the Resources page together with the story of how it was recovered): <http://www.archimuse.com/papers/nhprc/item-lvl.html>

Dublin Core metadata set current version (2011); review this originally resource-discovery oriented metadata set and also investigate how it was expanded as Qualified Dublin Core (the "terms namespace") at <http://dublincore.org/documents/dcmi-terms/>

Review the work being done for their own internal consistency by the Library of Congress Metadata for Digital Content Working Group here: <http://www.loc.gov/standards/mdc/index.html> See also the Master Metadata List, available on Canvas.

Investigate the PREMIS 2.2 metadata set for digital preservation and read the first 20 pages of this document: <http://www.loc.gov/standards/premis/v2/premis-2-2.pdf>



Investigate the METS metadata set for packaging digital objects and be prepared to discuss the parts of a METS document by reading the "METS Overview and Tutorial" (2011):  
<http://www.loc.gov/standards/mets/METSOverview.v2.html>

## **October 1: Passive vs active systems for managing desktop records**

### **New Technology presentation 3:**

**Topic:** Records Management Applications (RMAs) versus careful and systematic exploitation of existing software. Review the Department of Defense 5015.2 EDMS-RM model and commercial implementations of 5015.2-compliant RMAs, practical efforts at implementation in Texas, Kansas, and Mississippi, automated vs creator-assigned classification, Microsoft's nascent efforts to invade this profit space using features of its widely-used integrated business system SharePoint, and a suggestion on why much of this is doomed to failure without further study of how people manage their "own" records.

### **Questions to prepare for discussion:**

- 1) How might you be likely to be subjected to a digital records management application at work? If you have had such an experience, be prepared to tell us about it.
- 2) How detailed must a records management application be in order to actually manage records, all records? Examine the STD 5015.2 for ideas as to what it must cover.
- 3) Would you consider (have you considered) outsourcing your entire personal recordkeeping to Google or another cloud host? How would you set up such a thing, and what would you want to consider?

### ***Readings***

*PIM*, pp. 90-166, chapters 6-9.

DoD 5015.02 specifications (latest version, dated 2007). Although this standard was first developed for the Department of Defense, it is now standard for all Federal agencies. This is a big document, but I want you to look through it carefully so you can see the level of detail that a government standard includes, understand how the tasks are parsed out to different records managers, and understand what the federal standard proposes to be able to manage (on your behalf): <http://www.js.pentagon.mil/whs/directives/corres/pdf/501502std.pdf>

Here is a blog entry on problems with DoD 5015.02 and its deployment by an expert, Don Lueders (be sure to check out the linked replies from others who don't agree):  
<http://www.aiim.org/community/blogs/expert/On-Why-I-No-Longer-Support-the-DoD-50152-Standard>



And here is a report of huge amounts of DoD data erased during the wars in the Middle East from operational laptops: <http://www.propublica.org/article/lost-to-history-missing-war-records-complicate-benefit-claims-by-veterans>

NARA, Continuing Study of Federal Agency Recordkeeping Technologies (2008), a report from NARA on how well the application of the standard outlined in STD 5015.2 is going in federal agencies: <http://www.archives.gov/records-mgmt/resources/recordkeeping-tech-2008.pdf>

Patricia Galloway, "Big Buckets or Big Ideas: Classification vs Innovation on the Enterprise 2.0 Desktop," (2008). This paper outlines the so-called "Big Buckets" approach to making desktop records management easier to use and questions its blanket usefulness for records that may be among the most important to keep, available here: <http://www.armaedfoundation.org/pdfs/BBpaper30.pdf>

Barry Wheeler, "Personal Archiving--Year End Boot Camp," 1/20/2012 entry about real-life personal digital archiving in the Library of Congress-sponsored blog *The Signal: Digital Preservation*, by a Digital Projects Coordinator at the LoC: <http://blogs.loc.gov/digitalpreservation/2012/01/personal-archiving-%E2%80%93-year-end-boot-camp/>

## **October 8: Genres of digital records and their management**

### **New Technology presentation 4:**

**Topic:** Genres of digital records that lack paper analogs and their characteristics and problems. Review of desktop applications output, email, SMS/IM, websites/blogs/wikis, databases, still images, audio and video, etc. Note that many of these genres, especially (but not exclusively) when they are owned by individuals, are migrating into the cloud or never lived anywhere else.

### **Questions to prepare for discussion:**

- 1) Review these categories of digital objects in terms of your own personal information plan: which of these do you have? Where are they? What do they mean to you?
- 2) How important are format standards here? What are format standards for? Do you know what the formats of all of your nontext holdings are?

**Readings:** *Sample among these resources, picking at least one under each bold heading to talk about in class. We will talk about all of these in this class and the next.*

### **General:**

*Building a National Strategy for Preservation: Issues in Digital Media Archiving* (CLIR, April 2002) provides a series of short summaries of the problems of different genres and media: <http://www.clir.org/pubs/reports/pub106/contents.html>



Abby Smith, "Distributed Preservation in a National Context," D-Lib Magazine, Vol. 12, No. 6, June 2006, available at: <http://www.dlib.org/dlib/june06/smith/06smith.html>. Four years later, the Library of Congress reported on the progress of its NDIIPP project to carry out preservation of mostly non-text materials.

### **Desktop documents:**

Devin Becker and Collier Nogue, "Saving-Over, Over-Saving, and the Future Mess of Writers' Digital Archives," *American Archivist* 75 (Fall/Winter 2012), 482-513. Available from PCL.

### **Email (and other messaging):**

United Nations, Managing emails as records (2006):

<http://archives.un.org/unarms/en/unrecordsmgmt/electronicrecords/manageemailrecs.html>

Marlan Green, Sue Soy, Stan Gunn, Patricia Galloway, "Coming to TERM: Designing the Texas Email Repository Model," *D-Lib*, September 2002:

<http://www.dlib.org/dlib/september02/galloway/09galloway.html>

Chris Prom (archivist, University of Illinois at Urbana-Champaign) blog entries on email archiving: <http://e-records.chrisprom.com/recommendations/develop-submissioningest-policies/email-management-and-preservation-advice/>

### **Websites, blogs, wikis, social media:**

Look at the development of ideas and concepts in web archiving from 2001 to 2010 using the annual WebArchiving workshops, which are all linked to from this page:

<http://iwaw.europarchive.org/> The latest proceedings (of IWAW 2010; after 2010 this workshop was folded into the iPRES conference) are here: <http://www.iwaw.net/10/IWAW2010.pdf>

Marieke Guy, "Case Study: Archiving the JISC PoWR blog" (2010) <http://blogs.ukoln.ac.uk/jisc-bgdp/2010/10/11/case-study-archiving-the-jisc-powr-blog/>

"Archiving Social Networking Sites with Archive-It" (2011; aimed at Facebook--but note that for the time being Facebook has stopped this--as well as Twitter, YouTube. It is basically screen-scraping; can you think of how if you were FaceBook you would be archiving everything from the "inside"?):

<https://webarchive.jira.com/wiki/display/ARIH/Archiving+Social+Networking+Sites+with+Archive-It>

### **Databases:**

DPE Briefing Paper, Database Preservation: The International Challenge and the Swiss Solution: [http://www.digitalpreservationeurope.eu/publications/briefs/database\\_preservation.pdf](http://www.digitalpreservationeurope.eu/publications/briefs/database_preservation.pdf)

### **Still Images**



Cambridge in Colour website, "Archival Digital Photo Backup" available at <http://www.cambridgeincolour.com/tutorials/archival-photo-backup.htm>

Kristina Spurgin, "'Three Backups Is a Minimum': A First Look at Norms and Practices in the Digital Photo Collections of Serious Photographers," in *I, Digital: Personal Collections in the Digital Era*, ed. Christopher Lee (Chicago, SAA, 2011), 151-201. Available on Canvas.

### **Audio, Video, "New Media":**

Academy of Motion Picture Arts and Sciences, *The Digital Dilemma: Strategic Issues in Archiving and Accessing Digital Motion Picture Materials* (2007), an industry viewpoint, available here: <http://www.oscars.org/science-technology/council/projects/digitaldilemma/>

Academy of Motion Picture Arts and Sciences, *The Digital Dilemma II: perspectives from Independent Filmmakers, Documentarians, and Nonprofit Audiovisual Archives* (2012)--brand new, some additional perspectives: <http://www.oscars.org/science-technology/council/projects/digitaldilemma2/index.html>

Nan Rubin, *Preserving Digital Public Television* (2010), <http://www.thirteen.org/ptvdigitalarchive/files/2010/07/FinalReport.pdf>

Piotr Adamczyk, "Ethnographic Methods and New Media Preservation," presented at Museums and the Web 2008: <http://www.archimuse.com/mw2008/papers/adamczyk/adamczyk.html>

## **October 15: Centralized vs distributed models: custodianship**

### **New Technology presentation 5:**

**Topic:** Where digital records should be archived and by whom. Discuss the issue of traditional archival custodianship, the challenge of postcustodial models, and the emergence of best practice in the form of the OAIS repository model.

### **Questions to prepare for discussion:**

- 1) Can digital archives be "a place"?
- 2) Should there be a distinction between public and private records?
- 3) Should public records preservation be outsourced? Why or why not?
- 4) What would the individual person's point of view be on custodianship? What cloud locations might individuals use?

### **Readings**



Luciana Duranti, "Archives as a Place," *Archives and Manuscripts* 24(2): 242-255 (1996). Available on Canvas.

OCLC-RLG, "Trusted Digital Repositories: Attributes and Responsibilities," May 2002, AKA "OAIS Lite," available at: <http://www.oclc.org/research/activities/past/rlg/trustedrep/default.htm>

Library of Congress, "How to preserve your own digital materials," <http://www.digitalpreservation.gov/you/>

## **October 22: Maintaining the archival bond: Provenance and context**

### **New Technology presentation 6:**

**Topic:** Provenance and how to maintain it. Discuss what provenance is and how provenance can be provided for digital records; discuss the complexities of multiple or joint provenance issues and changes/accumulation of provenance history over time.

### **Questions to prepare for discussion:**

1) How can you establish the provenance for records that you create? Experiment with this: just look at a file in one of your directories and then see what properties you can see about it (in Windows you'll at least be able to see stuff like when it was created); now open it in your word processor and look at properties again--you should see some additional information. Where is this information coming from? How accurate is it?

2) Go back to the 5015.02 requirements and the discussion we had around it; how does the 5015.2 STD propose to build in maintenance of provenance?

### ***Readings***

David Bearman and Richard Lytle, "The Power of the Principle of Provenance," from *American Archival Studies: Readings in Theory and Practice*, 2000, 345-360, available on Canvas.

Maria Guercio, "Archival Theory and the Principle of Provenance for Current Records," from *The Principle of Provenance*, 1994, 75-86, available on Canvas.

Raimo Pohjola, "The Principle of Provenance and the Arrangement of Records/Archives," from *The Principle of Provenance*, 1994, 87-101, available on Canvas.

## **October 29: Permanence: media, formats, migration, emulation**

### **New Technology presentation 7:**



**Topic:** How to preserve digital objects over time. We'll discuss two important aspects: what "digital preservation" means and what it is we are trying to preserve.

**Questions to prepare for discussion:**

- 1) Considering your personal records, what would you think of as "good enough" preservation for text? What about for photographs? (You can also refer to readings we have already discussed.)
- 2) What are the major obstacles that you have yourself seen to keeping digital objects that you have created over time?

**Readings**

Caroline Arms and Carl Fleischhauer, Sustainability of Digital Formats: Planning for Library of Congress Collections: <http://www.digitalpreservation.gov/formats/intro/intro.shtml>

Jeff Rothenberg, "Ensuring the Longevity of Digital Information," (1999) available at: <http://www.clir.org/pubs/archives/ensuring.pdf>

This is the serious advocacy piece about archival emulation and Rothenberg has continued to support it until emulation has become more and more important.

Maureen Potter, " Researching Long Term Digital Preservation Approaches in the Digital Preservation Testbed (Dutch Testbed Digitale Bewaring)," *RLG Diginews*, June 2002, available at: <http://worldcat.org/arcviewer/2/OCC/2009/08/11/H1250008792610/viewer/file2.html>

Phil Mellor, Paul Wheatley, and Derek Sergeant, "Migration on Request, a Practical Technique for Preservation" in *Lecture Notes in Computer Science* (Springer, 2002), 516-526. This piece shows the simple argument for why it is crazy to use the chain-of-interpreters form of migration. Available at <http://www.springerlink.com/content/752vmvw0g0w40dj2/> If you can't get this without paying, go through the library catalog.

**November 5: Guaranteeing authenticity: security vs access**

**New Technology presentation 8:**

**Topic:** Authenticity vs access. Discuss the requirements of security for the preservation of digital records.

**Questions to prepare for discussion:**

- 1) How can you make sure that a digital object has not been changed? How likely is it that visual inspection of any kind would be adequate?



2) What is an "authentic" digital object? How can a digital object be more or less authentic? Is this a black-and-white issue?

### ***Readings***

Peter Hirtle, "Archival Authenticity in a Digital Age," *Authenticity in a Digital Environment* (Washington: CLIR, 2000), 8-23; available at: <http://www.clir.org/pubs/reports/pub92/hirtle.html>  
The whole report is well worth reading for an overview, since the issues have not changed.

InterPARES, "Findings on the Preservation of Authentic Electronic Records," September 2002; this is the set of principles that the National Archives is using, for better or for worse (focus on pages 11-21), was available at: <http://www.gseis.ucla.edu/us-interpares/pdf/InterPARES1FinalReport.pdf>; as an exercise, find this by shoving the URL into the Wayback Machine.

Borja Sotomayor, *The Globus Toolkit 4 Programmer's Tutorial* (2009), Chapter 9: "Fundamental Security Concepts" (do all the sections of Chapter 9), available at: <http://gdp.globus.org/gt4-tutorial/multiplehtml/ch09.html>

### **November 12: Experimentation with evaluation tools**

**Topic:** For today's class TA Pearl Ko will lead the class in an applied lab using digital tools to review and evaluate email.

### **November 19: Dealing with ownership: Gating vs sharing**

#### **Personal Electronic Records Management Plan DUE**

**Topic:** Discussion of intellectual property issues in providing access to digital records. Also look at the issues raised by information that others own (including public information) available all over the Web for people to aggregate and sell..

#### **Questions to prepare for discussion:**

- 1) It's pretty easy to copy a digital object and use it for anything you want. For example, what do you think about music sharing and the reaction of the music industry?
- 2) And have you ever heard the expression "Information wants to be free"? What does that mean? How expensive is it to reproduce digital information?
- 3) Looking at the American copyright law document's listing of enactments, notice how new technologies affect the enactment of new law.



4) If you had written something of which you were very proud and wanted to share it with others, which of the Creative Commons licenses might you choose to protect it? Why did you choose it?

5) What might be the status of personally significant records/information/data that you don't physically control?

### ***Readings***

Lessig, *The Future of Ideas*, Chapter 6, "Commons Lessons," available on Canvas; and look at the Creative Commons website: familiarize yourself with what a CC license is and the kinds of them there can be. <http://creativecommons.org/>

Current U.S. copyright law, Circular 92: <http://www.copyright.gov/title17/> This is a huge document., Look at the "statutory enactments" listed in the **preface** and then examine the **appendices** referring to the major versions including 1976 and following in the appendices.

Peter Suber, "Open Access Overview" (2004).  
<http://www.earlham.edu/~peters/fos/overview.htm>

World Economic Forum (the folks from Davos), *Personal Data: The Emergence of a New Asset Class* (2011), is just one example of how interested others might be in your personal data, available here:  
[http://www3.weforum.org/docs/WEF\\_ITTC\\_PersonalDataNewAsset\\_Report\\_2011.pdf](http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf)

If you own a home, look it up by typing your address into Google and then see what some of the real estate sites know about you from public databases. If not, look yourself up on Spokeo.

### **November 26: Access and markup: finding aids, internal markup, metadata, and search**

#### **New Technology Presentation:**

**Topic:** Markup: what it is and what kinds are most important. Discuss markup as a resource discovery aid and especially the level of granularity of markup.

#### **Questions to prepare for discussion:**

- 1) How is it useful to embed tags into text? What kind of embedded tags do you use every day?
- 2) How are tags used on webpages to assist in search?
- 3) How is EAD markup related to digital objects kept in archives?
- 4) Will conventional finding aids to archival collections become obsolete? What does the literature tell us about how easy (or not) they are to use for different audiences?



5) What kind of value added does an archivist bring to a fonds by creating an archival finding aid?

### ***Readings***

Text Encoding Initiative (TEI) reference page: <http://www.tei-c.org/Support/Learn/intro.xml>

Anne Gilliland-Swetland, "Popularizing the Finding Aid: Exploiting the EAD to Enhance Online Discovery and Retrieval in Archival Information Systems by Diverse User Groups," in Pitti and Duff (eds.), *Encoded Archival Description on the Internet*, 199-225 (2001). Available on Canvas.

Ian Witten, "Text Mining," in *Practical Handbook of Internet Computing*, 2005, ed. M.P. Singh, pp. 14-22: <http://www.cs.waikato.ac.nz/~ihw/papers/04-IHW-Textmining.pdf>

Marieke Guy and Emma Tonkin, "Folksonomies: Tidying up tags?," D-Lib 12(1), January 2006. <http://www.dlib.org/dlib/january06/guy/01guy.html>

Beth Yakel and Polly Reynolds, "The Next Generation Finding Aid..." Case study from New Skills for a Digital Era workshop, June 2006: <http://www.archivists.org/publications/proceedings/NewSkillsForADigitalEra.pdf>

Mary Flanagan and Peter Carini, "How Games can Help Us Access and Understand Archival Images," *American Archivist* 75 (Fall/Winter 2012), 514-537.

To expand on this article, view the talk by Luis von Ahn here: <http://www.cs.cmu.edu/~biglou/> (click on the item labelled "Google tech talk" to see the Human Computation video).

### **December 3: Business Plan Presentations**

**Topic:** Student teams will present their business plan projects to the class and the class will offer comments on the presentations.

**Digital Preservation Business Case Planning Project report DUE**



## Other Resources

Inventory document to use in preparation of personal records plan:

<http://courses.ischool.utexas.edu/galloway/2011/spring/INF389G/Inventory2011.doc>

NARA Email presentation materials: [Bibliographic sources](#), [slideset](#)

Pittsburgh Project (digital recovery) website: <http://www.sis.pitt.edu/~bcallery/pgh/index.htm>

Several British projects on personal recordkeeping:

The Paradigm project: Workbook on digital private papers:

<http://www.paradigm.ac.uk/workbook/index.html>

The Versions toolkit: [http://www.lse.ac.uk/library/versions/VERSIONS\\_Toolkit\\_v1\\_final.pdf](http://www.lse.ac.uk/library/versions/VERSIONS_Toolkit_v1_final.pdf)

NHPRC project links:

<http://www.archives.gov/nhprc/projects/>

InterPARES bibliography of reports from NHPRC projects:

<http://www.gseis.ucla.edu/us-inter pares/bibliography/NHPRC.htm>

Creator Guidelines for individuals:

[http://www.inter pares.org/display\\_file.cfm?doc=ip2\(pub\)creator\\_guidelines\\_booklet.pdf](http://www.inter pares.org/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf)

PADI portal site for digital preservation information:

<http://www.nla.gov.au/padi/index.html>

## Resources on Digital Art:

Archiving the Avant Garde: Documenting and Preserving Variable Media Art:

[http://www.bampfa.berkeley.edu/about\\_bampfa/avantgarde.html](http://www.bampfa.berkeley.edu/about_bampfa/avantgarde.html) There are many important links on this site, including an important report from the Guggenheim, Alain Depocas, Jon Ippolito, and Caitlin Jones (eds.), *Permanence Through Change: The Variable Media Approach*, available at: <http://www.variablemedia.net/pdf/Permanence.pdf>

404 Object Not Found project

[http://www.404project.net/hintergruende/das\\_projekt/index\\_e.html](http://www.404project.net/hintergruende/das_projekt/index_e.html)

## Preservation project resources:

Dutch digital preservation testbed result papers (note there are three major reports on migration, emulation, and XML as a preservation strategy):

<http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185&categorie=2>

Public Citizen's online collection of documents on NARA's e-records management direction (with links to many significant reports):

<http://www.citizen.org/litigation/briefs/ERecords/articles.cfm?ID=10536>



Audio, Video, and Motion Picture Preservation resource pages, including many resources for preserving digital objects, compiled by alumna Hannah Frost, on Conservation On-Line (CoOL): <http://cool.conservation-us.org/>

**Interesting stuff:**

Presentation by Rick Prelinger titled "Are Archives Doomed?" that argues for opening archives to greater public access, delivered at the University of Pittsburgh on January 26 (click on the link on this page; the talk goes on for a little over an hour):

<http://mediasite.ciddle.pitt.edu/mediasite/viewer/?cid=0561474e-e57d-4c4d-b8e0-2b6c42bd700c>

**Miscellaneous Resources**

[UT Library Online](#)