INF 389J: Appraisal and Selection of Records, Fall 2012
Prof. Patricia Galloway

Assignment:

Formal Research Paper: Each student will write a formal paper on appraisal, 15-20 pages double-spaced and exclusive of endnotes, appendices, and bibliography. The paper should be a scholarly and professional production adhering to the most recent Chicago Manual of Style (because if you want to publish where most of the appraisal literature is published, that's what you'll need to use).

Now what will the paper be about? I have posted several sample topics below and students may review the syllabus and propose others for approval if necessary. Whatever topic you choose, you must then choose a local archives (i.e., an archives in Austin) to serve as the central example for your paper (you can use others with which you are familiar for ancillary examples). There are no exceptions to this requirement, because it is important for you to be able to experience the archives in its context in order to understand its appraisal practices. In carrying out your research, you are not to pester the local archivists, but investigate for yourself what they do, seeking out relevant documents, making yourself a connoisseur of finding aids and catalogs, and generally learning how to see an archives critically using only the information available to archives users and the community in which the archives is situated. This requirement may limit the local archives you can choose for your paper, but it may also stretch your ingenuity a little.

Potential Paper topics:

Each of these topic sentences designates a broader area than your paper will need to cover. I expect that you will come up with an idea in this area and will create for me a proposal in the form of a paragraph that gives me something of a flavor of the research you want to do, the direction you want to take, and the kind of arguments you might make.
Scientific history and the requirement to have archives
Government legitimacy and the requirement to have archives
Archival appraisal/reappraisal and its effects on cultural memory
Jenkinson and creator selection of materials to preserve
Schellenberg and value appraisal as selection criteria
Community participation in documentation strategy
Reflecting organizations through functional analysis/macro-appraisal
Appraise and/or schedule? Appraising public records
Appraisal policies for faculty papers
Monetary worth, desirability, and the appraisal of private records
Economics of archives as a cultural good in the long term
Effects on appraisal of the comparative costs of paper vs digital collections
The influence of use and customer base on appraisal: history and theory
How does appraising digital-only objects--databases, email, blogs, wikis—differ from appraising physical objects (or not)?
What if we Kept It All? (Discuss and contextualize the arguments pro and con.)

(source: http://courses.ischool.utexas.edu/galloway/2012/fall/INF389J/assignment.html)

Arcadia Falcone
December 3, 2012
INF 389J: Appraisal and Selection of Records
Final Paper

Digital Materials, Paper Practices, and the Appraisal of Literary Manuscripts at the Harry Ransom Center

Introduction

Appraisal theories and practices have adapted over time to take account not only of broader trends in archival theory and modes of scholarship, but also technological shifts. As the personal computer has supplanted the pen and the typewriter, and the Internet has invaded the domain of the telephone and fax, many archival institutions have extended their appraisal and collecting policies to encompass digital materials.[1] Manuscript repositories that collect archival artifacts produced by entities external to their organization have faced particular challenges in responding to the idiosyncrasies of how individuals have managed (or failed to manage) their personal digital archives, further exacerbated by the technical difficulties arising from the acquisition of digital files that may have been created many years prior to their arrival at the archive. These files may reside on legacy digital media that must be appropriately *mediated* by hardware and software before its data is even available as an object of appraisal. When the digital collections comprise records of literary or other creative activities, aesthetic considerations may mean that accessing the informational value of the content is inadequate for responsible appraisal. While digital archivists may approach some of these issues through analogy with paper manuscript collections, the particular affordances and lacunae emerging from digital formats at times require experimenting with new strategies. In Part I of this paper, I will

[1] While "digital materials" may seem an oxymoron, the phrase is also revealing: digital files ultimately reside on some kind of material medium, as the collections discussed in this paper make manifest.

examine the history of this experimental process at the Harry Ransom Center, focusing on how manuscript policies have and have not carried over into digital appraisal practices. In Part II, I will place the Center's digital appraisal decisions in the context of research projects aimed at understanding and conceptualizing appraisal principles for born-digital personal papers, evaluating the Center's acts of appraisal in relation to the extensive documentation of guidelines for best practices produced by Great Britain's Paradigm Project. I will conclude by considering how researchers' use of digital literary manuscripts reflects on current appraisal practices and indicates avenues for future exploration. While traditional manuscript policies have provided a useful basis for digital appraisal strategies in collecting archives such as the Ransom Center, analogies to paper too narrowly circumscribe the distinct and dynamic affordances of digital media.

Part I: Appraising Born-Digital Collections at the Harry Ransom Center

As of Spring 2010 (the most recent information I could locate on the Ransom Center website), the Center holds 39 "hybrid" collections comprising both paper documents and physical media with born-digital content. These include six collections processed by iSchool students in Dr. Patricia Galloway's INF392K course on digital preservation from 2005 to 2009: the Michael Joyce Papers (2005), the Arnold Wesker Papers (2007), the John Crowley Papers (2007), the Leon Uris Papers (2007), the Norman Mailer Papers (2008), and the Terrence McNally Papers (2009). As part of these course projects, students ingested collection materials, metadata, and documentation into the iSchool's DSpace digital repository. This documentation includes accounts of the appraisal decisions taken in consultation with the Center's digital archivist, Gabriela Redwine. The documentation (with the exception of the Mailer project, for

which the report is password-protected) is publicly accessible and appears in web searches, while the collection materials themselves are restricted, and available for viewing only on a dedicated laptop in the Ransom Center reading room. These project reports summarize the ongoing development and implementation of the Center's appraisal policies and practices for born-digital materials over several years.[2]

While each project group approached processing the born-digital files somewhat differently, the basic sequence of actions was similar across all groups. After examining and cataloging the physical media, the group members would copy the files to the hard drive of their working computer (at times a challenging task) on site at the Ransom Center, maintaining the original file directory order. Each group examined most if not all of the files individually, describing files at item-level except in a few special cases. Generally it was only after all files had been accessed and described that appraisal decisions would take place. Working copies were derived from the files recovered from the physical media, which were then in some cases converted to non-proprietary file formats or re-named with standard file extensions. The files would then be arranged, ingested into DSpace, and described in the existing finding aid for the paper portion of the collection.

Appraisal decisions for born-digital materials were largely based on the Ransom Center's existing practices for manuscripts, though some divergences developed over time. The Joyce project group identifies three main types of materials for disposal (which was usually

[2] In addition to these reports, I also consulted the other project documentation available on DSpace. The finding aids for these collections, when available, offered some additional information, but tended not to include much detail about electronic resources. An article from the Ransom Center's newsletter, available on their website, provided a sense of the larger context for these projects: Gabriela Redwine, "Preserving Born-Digital Materials," *Ransom Edition* (Spring 2010), http://www.hrc.utexas.edu/ransomedition/2010/spring/borndigital.html. Additional perspectives on the Michael Joyce project, though with less detail as to the process of appraisal, appeared in two additional articles from the project group members: Catherine Stollar and Thomas Kiehne, "Guarding the Guards: Archiving the Electronic Records of Hypertext Author Michael Joyce," in *New Skills for a Digital Era*, ed. Richard Pearce-Moses (Chicago, Society of American Archivists, 2008), 57-64; and Catherine Stollar Peters, "When Not All Papers Are Paper: A Case Study in Digital Archivy," *Provenance* XXIV (2006): 23-35.

accomplished across these projects through omission from the preservation and description facets of processing; the bitstreams on the accessioned physical media remained intact). The first category of materials the Joyce group slated for disposal is exact duplicates, as determined by identical MD5 hashes as well as file dates, formats, and sizes; this approach treats duplicate files as similar to carbon copies or photocopies, which are often removed from manuscript collections when the original is present. Derivative copies, when they add no discernible evidential or information value, are in both cases judged unworthy of archival preservation. The group report does, however, point out that "the files saved on Joyce's disks reflect his own appraisal": "Since the files on the disks we received were not only created by Joyce and saved to the hard drive, but they were saved again, if not multiple times as backups, by Joyce, we are preserving only those files Joyce intended to preserve."[3] While this interpretation of Joyce's intentions is open to debate, this perspective does suggest that Joyce's decision to save multiple backups (explicitly labeled by him as such) could provide evidence as to his creative practices and priorities, even if the files are themselves identical. Recognizing this principle guides the group processing the McNally collection in 2009 to take a different approach:

> In previous collection uploads for INF 392K, the working groups have eliminated duplicates to save time and storage space, avoid user confusion and upload problems arising from duplicate filenames, and streamline the overall ingest process. In the case of the McNally collection, these were not compelling arguments for deleting files that, if left in place as they were stored on the disks by McNally himself, could provide some insight into his habits of writing and computer use.[4]

While duplicates were retained in the McNally DSpace digital collection, which followed the file arrangement on the disks, the description of digital objects in the finding aid continued to follow

[3] Thomas Kiehne, Vivian Spoliansky, and Catherine Stollar, "From Floppies to Repository: A Transition of Bits" (course project report, UT Austin, 2005), 10. Available online at https://pacer.ischool.utexas.edu/handle/2081/941.
[4] Snowden Becker, Virginia Luehrsen, and Sarah Winblatt, "Bits Together, PINs Apart: Preparing the Terrence McNally Digital Records for DSpace Ingest and Research Access," (course project report, UT Austin, 2009), 10. Available online at https://pacer.ischool.utexas.edu/handle/2081/20185.

previous practices in arranging by series and omitting duplicates, so as "effectively [to] 'discard' redundant identical files […] simply by not mapping those duplicate files" to a series.[5] Paper manuscript practices thus still dominated in terms of how metadata for born-digital materials was inserted into existing description, as a reflection of appraisal decisions.

Besides duplicates, the other two categories of materials the Joyce project group appraised as appropriate for disposal were software application files (with the exception of Storyspace) and student works. The group also considered excluding all works by third parties, but ultimately decided to retain these materials in accordance with Ransom Center policy for paper collections, which also proscribed keeping student papers. In explaining why most software application files were disposed, the group invokes *respect des fonds,* suggesting that the applications are not "the result of an organic process reflecting the functions of the creator."[6] Joyce was, however, co-creator with Jay David Bolter and John B. Smith of the Storyspace hypertext writing environment, and so those software files were considered part of Joyce's *fonds* and therefore retained. While the Ransom Center in general collects artifacts of a creator's workspace, counting many literary figures' desks and typewriters among its holdings, the focus in these appraisal decisions is on the digital content Joyce created, rather than the digital environment which mediated Joyce's interactions with that content.[7]

Storyspace is not only Joyce's creation, but is also the software for which he composed and distributed his best-known work, the hypertext fiction *Afternoon, a story* (privately distributed from 1987, published by Eastgate Systems from 1990). Before the World Wide Web, Storyspace enabled a creator to construct a series of pages, or "nodes," connected by hyperlinks.

[5] Becker et al., 10.
[6] Richard Pearce-Moses, "A Glossary of Archival and Records Terminology," *Society of American Archivists*, http://www2.archivists.org/glossary.
[7] I am purposely setting aside the host of legal issues in preserving and emulating proprietary software, as doing that topic justice is beyond the scope of this essay.

In *Afternoon,* a narrative emerges as the reader navigates from one node to another by clicking on unmarked linked words within each node. Moving through the story-network in this interactive fashion, as enabled by the Storyspace software, is thus integral to experiencing the work, as much as reading the pages of *Middlemarch* in the correct order would be. In addition to the program's place in Joyce's *fonds* by virtue of his role in its development, preserving the software's affordances is thus effectively necessary to preserve the work, in terms of allowing continued access to the story-experience.

This distinction, between preserving Storyspace as a work that Joyce created, and maintaining it or an equivalent as a tool for preserving access to *Afternoon*, emerges in how the project group fulfilled both these goals. The Storyspace files that formed part of Joyce's collection as accessioned, and which he used for composing and accessing the first edition of the story, were preserved with the rest of the born-digital materials. But, rather than installing this version of the software on the laptop used for reading room access to the Joyce digital collection, a later version of Storyspace, with similar functionality but a different look and feel, was made available. The project report discusses this decision:

> Initially, the HRC was not very concerned about keeping the look and feel of the original files as the priority was on retaining the information. As the project progressed it was further discussed that, especially in the case of the hypertext novels where the aesthetics plays a major role, the importance of retaining as much as possible of the "look and feel" is of concern. We had access to the new version of Storyspace, so the hypertext novels at the moment are accessed with the limitations of the available version and technology. […] Emulation was considered to recover "look and feel," but disregarded because the files could be accessed with the new version of Storyspace.[8]

The Center's approach suggests a three-tiered hierarchy of priorities, each of which requires a different level of preservation and access. The first level, the Center's original position, focuses on the informational content, as in the text and structure that constitute the entity of the story as

[8] Kiehne et al., 16.

an intellectual object. A printout of the story text with indications of how the nodes connect would adequately preserve its basic informational value. The second level, as was eventually implemented, also considers functionality, preserving access to the experience of navigating from node to node within the story. Any software that can render the text, interpret the links, and allow the user to click through the narrative as originally designed would be sufficient to preserve the work as a functional object. Finally, conceiving the work as having aesthetic value visually as well as textually requires at least emulating the original software to reproduce its visual experience as closely as possible.

The Joyce project report cites Smith et al.[9] in describing two levels of digital preservation: "bit preservation," where the original bitstream is preserved, and "functional preservation," where "usability" is preserved. According to the group's conclusions, providing access to Joyce's Storyspace novels so as to reproduce functionality but not appearance fulfills the requirements of functional preservation, as the work is made similarly "usable": so long as the work is accessible, reproducing its look and feel is a secondary and ultimately disregarded concern. But should "usability" be the primary criteria for evaluating the authenticity of a particular representation an artistic work? To address this question, it is necessarily to contextualize it in broader issues of digital preservation and appraisal. At this point in time, digital archivists have thoroughly made the case for preserving the original bitstream as best practice; the work of appraisal is often to determine which bitstreams to keep. Unlike most paper manuscripts, however, preserving the object does not automatically entail preserving access to the object in a fundamental sense. While in most cases a paper document may be immediately apprehended by the human senses, digital media requires additional mediation

[9] MacKenzie Smith et al., "DSpace: An Open Source Dynamic Digital Repository," *D-Lib Magazine* 9 (2003), http://www.dlib.org/dlib/january03/smith/01smith.html.

through software and hardware for its information content to become available to human perception.[10] Thus the appraisal of digital objects must take into account not only the object to be preserved as a series of bits, but also what else needs to be maintained in order to translate those bits into something apprehensible by human senses. While the majority of paper documents offer combined content and access in a self-contained artifact, digital files require mediation to be accessible.

One task of digital appraisal, then, is to judge to what extent, and in what ways, the form this mediation takes is an integral part the digital object being preserved. This importance ranges from the basic technical functionality of taking the bitstream input and accurately transforming it into human-legible output, to full-fledged emulation or recreation of the original system hardware and software, with most instances falling somewhere between the two. Deciding where on the spectrum a particular object, collection, or format falls entails considering characteristics similar to what Margaret Hedstrom and Christopher A. Lee call "significant properties": "those properties of digital objects that affect their quality, usability, rendering, and behavior."[11] More precisely, these properties depend on the interaction between the digital object and the mediating hardware/software that translates characteristics of the digital object into perceptible attributes. The importance of this dependence varies according to how much the mode of mediation participates in creating these attributes or significant properties. For a digital document whose primary value is informational, the particular mode of mediation matters less, so long as it represents the contents accurately. But for digital objects operating as works of art, the mode of mediation may be integral to the work beyond simple functionality.

[10] I am purposely avoiding "understanding" or "reading" here in order to invoke the more fundamental level of being able to perceive what is there to be read, regardless of being able to understand or even to parse it.
[11] Margaret Hedstrom and Christopher A. Lee, "Significant Properties of Digital Objects: Definitions, Applications, Implications," *Proceedings of the DLM-Forum 2002*, 218.

Which brings us back to Michael Joyce's *Afternoon*, and the question of what role the aesthetic properties of its originating interface should play in appraising its constituent elements for preservation and access. As quoted above, the Joyce project report aligns the Center's initial assessment with informational content and usability, defining accessibility to the text and its structure as adequate for preservation. The later version of Storyspace offers both kinds of accessibility, and therefore is considered sufficient for these purposes. But in addition to possibly damaging the work's artistic integrity by altering its appearance, this mode of access substantially elides the multiple permutations of work and software over time. No fewer than six editions of *Afternoon* were issued from 1987 to 1998, with at least half of those also appearing in slightly different versions for Macintosh OS and Microsoft Windows. The first five editions developed in tandem with changes both in Storyspace and the Macintosh operating system, as the fifth edition's colophon describes:

> This fifth edition of *Afternoon, a Story* reflects Eastgate Press changes in jacket copy and front matter only. The fourth edition involved Storyspace changes to suit Macintosh © Systems 6.07 and above. The third edition (first published by Eastgate in 1990) changed text windows and typefaces and made minor fixes of links and texts, all differing from the second edition (1989) which took advantage of certain changes to Storyspace to add new links and create a few new places. The first edition of *Afternoon* was distributed quite informally to a number of Storyspace beta-users and interested writers and scholars beginning in 1987.[12]

The sixth edition appeared online via a Java interface (and, oddly, in print) as part of the *Norton Anthology of Postmodern American Fiction*.[13] In 2012, *Afternoon* is still available (for Mac OS X and 32-bit Windows 7) from the Eastgate Press website, which also promises a forthcoming version for iPad.[14] The permutations become even more complex when considering the interplay

[12] Matthew G. Kirschenbaum, *Mechanisms: New Media and the Forensic Imagination* (Boston: Massachusetts Institute of Technology, 2008), 160.
[13] Cited in Matthew G. Kirschenbaum, "Editing the Interface: Textual Studies and First Generation Electronic Objects," *Text* 14 (2002): 28.
[14] http://www.eastgate.com/catalog/Afternoon.html.

between Joyce's digital "drafts" of *Afternoon* in its various states of vision and revision, the first two privately circulated editions, the subsequent published editions, and multiple versions of both *Afternoon* and Storyspace for different operating systems. In conceiving the *Afternoon* digital files as self-contained, property-bearing entities akin to manuscript drafts, rather than in conjunction with the various distinct forms of mediation through which they manifested over time, the Ransom Center elides significant aspects of *Afternoon*'s complex textual and digital history.

Afternoon may constitute a boundary case, but the challenges it poses to traditional appraisal practices for literary manuscripts have wide-ranging implications for adapting these practices to digital materials. Considering the structures of mediation at play in any human-legible rendering of a digital file as something to look *at*, not *through*, must inform the appraisal decisions that govern preservation and access to digital materials.

Part II: A Paradigmatic Evaluation

While digital personal papers have received less attention in the appraisal literature than government and organizational records, several recent projects have focused on systematically adapting principles of appraisal developed for paper manuscripts or institutional electronic records to meet the challenges of born-digital materials within collecting archival repositories. In 2008-2009, the Harry Ransom Center joined Emory University's Manuscripts, Archives, and Rare Books Library (MARBL) and University of Maryland's Maryland Institute for Technology (MITH) for a grant project funded by the National Endowment for the Humanities' Office of Digital Humanities, entitled "Approaches to Managing and Collecting Born-Digital Literary

Materials for Scholarly Use."[15] Project Director Matthew G. Kirschenbaum has written much in various forums about preserving digital manuscript materials from a literary scholar's perspective, including some issues related to appraisal. Michael Forstrom published his case study "Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library" in *The American Archivist* in 2009.[16] The 2011 collection *I, Digital: Personal Collections in the Digital Era* includes one essay devoted to appraisal by Christopher A. Lee, "Collecting the Externalized Me: Appraisal of Materials in the Social Web."[17] In addition, some work done by long-running digital preservation research projects, such as International Research on Permanent Authentic Records in Electronic Systems (InterPARES) and Creative Archiving at Michigan and Leeds Emulating the Old On the New (CAMiLEON) is not specific to digital personal papers but still offers applicable findings.

Perhaps the most comprehensive account so far of managing born-digital personal papers, however, is the Paradigm Project. This British project, spearheaded by the Bodleian Library at the University of Oxford and the John Rylands University Library at the University of Manchester, undertook research on managing digital personal papers from 2005 to 2007. The end result was the development of best-practice guidelines documented in *The Digital Private Papers Workbook*, published on the project website.[18] This workbook includes a section devoted specifically to appraisal and disposal, drawing upon principles of macroappraisal and functional appraisal, which provides an unusually detailed account of this aspect of archival management for digital personal papers. While some aspects of the project are specific to the British legal

[15] Matthew G. Kirschenbaum et al, "Approaches to Managing and Collecting Born-Digital Literary Materials for Scholarly Use," (White Paper to the NEH Office of Digital Humanities, Level 1 Digital Humanities Start-Up Grant, May 2009).
[16] Michael Forstrom, "Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library," *The American Archivist* 72 (2009): 460-477.
[17] Christopher A. Lee, ed, *I, Digital: Personal Collections in the Digital Era* (Chicago: Society of American Archivists, 2011.
[18] http://www.paradigm.ac.uk/workbook/.

context and the pilot's focus on politicians, it still provides useful criteria for examining the relation between Ransom Center's appraisal of digital personal papers and a set of best practices formulated based on principles of appraisal theory. While the Center's practices in many cases parallel the conclusions reached by the Paradigm Project, the project workbook suggests some avenues for further development towards strategies that take greater account of the differences between how digital and paper may be managed in personal archives.

In its section on appraisal, the Paradigm workbook asks the fundamental question, "Is it necessary to appraise digital records at all?" To address this query, the workbook provides a summary of arguments both pro and con. The arguments against appraisal include, in addition to the inexpensiveness of storage, the necessity of adapting to creators' own record-keeping practices: as many creators take advantage of increasingly refined search tools to locate individual files rather than organizing their records into hierarchical directories, the workbook argues, the archivist must either adopt a similar approach or undertake a prohibitively labor-intensive survey of the files in order to gain enough intellectual control even to begin the appraisal process. On the other side, arguments in favor of appraising digital records include the acknowledgement that not only storage but also processing and preservation do entail costs, and, with the large amount of digital data being constantly produced, keeping everything is an unnecessary drain on resources. Although this section presents both perspectives, overall the workbook treats appraisal as "the primary function on which all other functions depend," maintaining its ongoing importance while recognizing that how it is practiced will continue to change with digital records and record-keeping practices.

Each of the Ransom Center digital collections discussed in Part I underwent item-level appraisal. For most of the collections, this was necessary in order to gain intellectual control

over the materials, and to address the preservation issues of legacy media. While the Paradigm workbook recommends early communication about orderly digital record-keeping practices with the creator, the lack of consistent organization or metadata in the majority of these collections suggest that they reflect their creators' idiosyncratic approaches. The major exception is the Arnold Wesker collection, which was arranged into directories by subject and year, and which also included supplemental metadata. The report for that project notes that these creator provisions greatly aided in quickly determining which files should be kept and which could be disposed (primarily duplicates). The majority of these collections, however, appear to consist primarily of files saved to floppy disks as ad hoc backups, rather as a systematic record of documents within the context of their initial creation, therefore requiring examination to determine content, purpose, and authorship, among other attributes relevant to appraisal. All of these collections were limited enough in size that item-level appraisal was feasible, though some groups recommended adopting toolkits when possible to automate some file analysis tasks. For these collections, a detailed appraisal of their contents was appropriate, but the issues raised by the Paradigm Project will continue to be relevant as digital content becomes more prevalent and numerous in the literary manuscript collections that the Ransom Center acquires.

The Paradigm workbook identifies three primary considerations affecting the appraisal of personal archives. First, as previously discussed, the creator's personal record-keeping practices determine the degree and kind of intellectual control the archivist can assume over the files upon their accession, and thus influence the amount of preprocessing necessary to begin appraising the collection materials. The second factor concerns the length of the time period between ingesting digital records into the institution's digital repository and appraisal. The workbook recommends making appraisal decisions at the point of ingest when possible, to avoid the wasted labor of

creating technical metadata for files that will be subsequently disposed. This guideline assumes a workflow in which ingest occurs simultaneously with or soon after accession, which was not necessarily the case for the Ransom Center digital materials. The Michael Joyce collection was ingested to DSpace in 2005, the same year as its accession, but ingestion for the other digital collections did not occur until three to seven years after accession, with item-level appraisal being carried out at that point. As the physical media (floppy disks) were accessioned without being ingested to a local server or other digital repository until appraisal and processing, the workbook's concerns about wasted resources are less applicable, though in some cases floppy disks were cataloged whose contents were ultimately omitted from ingestion. The delay may, however, have negatively affected appraisal, in that the hardware and software needed to access the files on the disks was less available or accessible due to obsolescence, or the media itself may have become corrupted over time. For these collections, the paper portion was processed soon after their accession, meaning appraisal decisions for digital files drew on what had already been done with the manuscripts, rather than the entire collection being assessed as an organic whole. While the technical logistics of dealing with digital files explain the Ransom Center's policy of separating digital media from hybrid collections, the Paradigm workbook suggests that appraising both digital and paper files concurrently would be preferable, as they are part of the same *fonds*.

The workbook's third consideration for the appraisal of personal digital files is the particular digital collection development approach that the repository pursues. A more proactive approach would take snapshots of the creator's digital files periodically, in which case the archive may also be able to shape the creator's record-keeping practices to better support archival appraisal upon accession. The Ransom Center follows what the workbook describes as

a more "traditional" approach, in that the archive acquires digital files on physical media that are no longer being actively used by the creator. This approach renders appraisal more resource-heavy, to the extent that the recovery of some files may be technically possible but practically unfeasible. While in many cases literary manuscript collections are not obtained until after the creator's death, seeking to engage living creators in organizing or at least documenting the management of their digital records would facilitate appraising those records at or even before the point of accession, which could streamline their processing.

In conjunction with working to improve creators' record-keeping practices, the workbook also recommends surveying digital records prior to accession, in order to assess their content, context, structure, and technical state. Drawing upon the principles of macroappraisal, this approach may enable particular categories of records to be identified for disposal relatively easily. While the variable contents of the floppy disks in the collections suggests that the Ransom Center did not survey the files prior to their acquisition, the project groups did apply similar scoping principles on a case-by-case basis as part of their appraisal, for example not preserving software files found in the Leon Uris and Michael Joyce collections, and omitting student papers that were also among the Joyce materials. Continuing to develop these collection-based decisions into a set of comprehensive scoping guidelines generally applicable to the kinds of files found in digital literary manuscript collections, much as the Paradigm Project did for political papers, would be useful in shaping appraisal strategies whether applied before or after accession.

Of the five general appraisal recommendations offered by the Paradigm Project, the Ransom Center digital archives projects demonstrate four to some degree. Each project tailored its appraisal criteria to the particular collection; continuing to develop these principles for

general policy would benefit future appraisal activities. All but one of the projects used checksums to identify and eliminate duplicates; the one exception, for the Terrence McNally collection, specifically chose to keep the duplicates as evidence of McNally's record-keeping practices. All of the projects did not preserve standard operating system or application files, except in the case of Storyspace where the software was determined to constitute part of Michael Joyce's *fonds*. While the Arnold Wesker collection arrived with substantial contextual information for the digital files, which contributed to the appraisal process, other creators were less scrupulous about documenting context, which sometimes complicated the identification and evaluation of digital files. Additionally, the separation of digital media from paper manuscripts, processed at different times and by different people, made it more difficult to match relevant contextual information within the manuscripts to the digital files. The final recommendation, to encourage best practice at the start of the record cycle, is understandably difficult given the Ransom Center's usual circumstances for accessioning manuscript collections, particularly those in which paper rather than digital materials are the prime consideration driving a particular acquisition. As digital files become more dominant in new accessions, the Paradigm Project suggests that the Ransom Center may benefit from moving away from appraisal practices designed for paper collections, and shifting appraisal earlier in the acquisition process.

Conclusion: Access, Use, and the Ultimate Ends of Appraisal

In an article for *The Chronicle of Higher Education* entitled "Hamlet.doc? Literature in a Digital Age," Matthew Kirschenbaum recounts his experience working with the Michael Joyce digital collection in the Ransom Center's reading room:

> The first accession of virtual materials […] has been lifted from the almost 400 diskettes that make up their original storage media and uploaded to an electronic repository system

> known as DSpace. They are online but can be accessed only from a dedicated laptop located in the center's reading room. To actually work with the files, I had to download them to the desktop of the machine, where I used what means and know-how I could to get the cranky old binaries to execute on the up-to-date operating system. Sometimes I was unsuccessful. […] Since DSpace maintains the integrity of a master copy of every file, I could do what I pleased with the derivative I downloaded to my local desktop — hack at it, tweak it, break it.[19]

Appraisal seeks to select for preservation materials that have value, and one (though by no means the only) measure of value is through the uses made of archival resources, and, for digital collections that have been accessible for less than a decade, what current scholars conceive as potential directions for future research. While ultimately unreliable indicators of what the future will actually bring, scholarly imaginings as to what the affordances of digital manuscripts might make possible do at least offer a rough contour map of what researchers may demand from digital archives over the next decade. Considering these past, present, and future uses of digital literary collections may inform, though not define, appraisal decisions.

Earlier in the same article, Kirschenbaum elaborates on the potential information that a researcher might glean from the digital record by imagining what might be offered up by Shakespeare's hypothetical hard drive:

> We might be able to know, for example, the precise date on which he began composing Hamlet indeed the precise minute and hour, time-stamped to the second. We would be able to know how long he had spent working on it, or at least how long the file containing the play had remained open on his desktop. We would very likely have access to multiple versions and states of the file, and if Shakespeare had "track changes" turned on while he wrote, we would be able to follow the composition of a soliloquy keystroke by keystroke, each revision also date- and time-stamped to the second. We might discover the play had originally been called GreatDane.doc instead of Hamlet.doc. We might also be able to know what else he had been working on that same day, or what Internet content he had browsed the night before (since we'll assume Shakespeare had Web access too). While he was online, he might have updated his blog or tagged some images in his Flickr account, or perhaps edited a Wikipedia entry or two. He might even have spent some time

[19] Matthew Kirschenbaum, "Hamlet.doc? Literature in a Digital Age," *The Chronicle of Higher Education*, August 17, 2007, http://chronicle.com.ezproxy.lib.utexas.edu/article/Hamletdoc-Literature-in-a/6887/.

> interacting with others by performing with an avatar in Second Life, an online place where all the world is truly a shared virtual stage. [20]

As a thought experiment, this vision is a self-admittedly grandiose fantasy of plenitude. It does, however, offer an idea of the kinds of evidence future scholars may draw upon in analyzing digital works. Kirschenbaum highlights how the versatility of digital devices, in terms of both their uses and their self-documenting capabilities, may create a body of evidence previously unavailable. (What researchers decide to *do* with all this information, what arguments this evidence will support, to a large extent remains to be seen.) While here the emphasis is on the informational cornucopia produced by digital affordances, elsewhere Kirschenbaum has addressed one of what might be termed the *anti*-affordances of digital composition: the reconfiguration of the "draft" when each new revision overwrites its predecessor, which, depending on the system settings, may be more or less accessible for future review by scholars and even by the creator herself.[21] Kirschenbaum's current project, *Track Changes: Authorship, Archives, and Literary Culture After Word Processing*, examines the implications of this shift in the material conditions of composition.[22] This holistic literary-critical mode in which the creator's corpus is composed of the entirety of her digital activities, whether or not they clearly constitute part of her creative work, would require a model of appraisal in which preservation of a complete digital record is the end goal—though "completeness" is a far from simple proposition when elements of that record are strewn across multiple personal devices and cloud storage facilities.

The extent to which Kirschenbaum's vision will appear prescient or quaint a decade from now awaits the test of time. In the present moment, the scholarly uses of the Ransom Center's

[20] Kirschenbaum 2007.
[21] http://otal.umd.edu/~mgk/blog/archives/000848.html.
[22] https://mkirschenbaum.wordpress.com/2011/04/10/track-changes/.

digital collections remain largely unexplored. In the two years leading up to the NEH digital humanities grant project in 2009, only three users had accessed digital materials at the Ransom Center, from the Michael Joyce and Arnold Wesker collections.[23] While scholars such as Kirschenbaum and Terry Harpold refer to the Joyce collections in their excavations of *Afternoon*'s complex digital history,[24] searches on JSTOR and Google Scholar turned up few citations of the Ransom's digital collections; those that did appear focused more on the challenges and affordances of the digital medium itself, rather than on its evidentiary value for specific arguments. While use does not constitute the sole criterion for determining archival value, the impact of access to digital manuscripts on scholarly concerns will continue to contribute to the ongoing development of appraisal theory and practice.

[23] Kirschenbaum et al. 2009, 11.
[24] Kirschenbaum 2002; Kirschenbaum 2008; Terry Harpold, *Ex-Folations: Reading Machines and the Upgrade Path* (Minneapolis, University of Minnesota Press, 2009).