



# On Digital Archives: Lessons from the Susan Sontag Hard Drives

## Citation

Hyry, Tom. 2015. On Digital Archives: Lessons from the Susan Sontag Hard Drives. In Proceedings of the Society of American Archivists meeting, Acquisitions and Appraisal Section Annual Meeting, Cleveland, OH, August 20, 2015.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:40918991>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

## **On Digital Archives: Lessons from the Susan Sontag Hard Drives**

**Tom Hyry, Florence Fearington Librarian of Houghton Library, Harvard University**

Delivered at the Society of American Archivists Annual Meeting

This paper relates to work I was involved in when I was Director of UCLA Library Special Collections, to acquire two hard drives of documents, correspondence, email, photographs, and other materials created and compiled by Susan Sontag over the course of the last dozen years of her life. This process raised technical, ethical, philosophical, financial, and practical issues that still seem new to the archival endeavor. The paper features a narrative of this acquisition, intended as a case study that allows us to examine multiple dimensions in the changing nature of archives and archival institutions. In particular, it considers how born digital materials should and do factor into the overall marketplace for literary archives, which remains largely undetermined.

At its core, this is a discussion about value, in at least two senses of the word. For generations libraries and archives have acquired the papers of literary figures<sup>1</sup> to fulfill their cultural heritage, research, and educational missions. Institutions and their users value the papers of literary figures for what they contribute to the literary legacy of their creators and for what they convey about the creative process, relationships with other writers and figures, and the socio-cultural scene in which they lived and worked. In this sense, these collections can be said to be invaluable; using the professional term, they are archival, records that have ongoing value for, in this case, the historical and cultural information they contain which justifies their continued preservation.

The papers of literary figures often have a value more commonly understood: financial. Some archival institutions regularly acquire through purchase the papers of remarkable individuals and organizations. The papers of literary figures have enjoyed a particularly robust trade and there are numerous examples of their archives being acquired by institutions for dollar figures that can stretch into the seven figures and up.<sup>2</sup> These collections and their individual contents have value for private collectors as well.

The status of writers' papers as artifacts constitutes a major part of literary archival value. These materials intrinsically trace a direct line between a literary figure and an object. When one works with an original manuscript, a letter, or a diary, one gets a sense of privileged access to a writer's true thoughts, intentions, and processes. In an analog setting, one acquires documents that not only contain this privileged and unknown information but that have a physicality to them that can be traced back to their creator. One holds the note Aldous Huxley wrote on his death bed; analyzes Virginia Woolf's handwriting from her own ink; delights in a

---

<sup>1</sup> In this article, literary archives stand in for all contemporary archival material for which there exists a market value. Papers of artists, performers, and other public figures often are offered for sale to individuals and institutions and the same issues raised in this article apply to these types of collections as well.

<sup>2</sup>A particularly readable entrée into the world of literary archives and their acquisition can be found in D.T. Max's profile of Thomas Staley in the *New Yorker*: Max, D.T. "Final Destination," *The New Yorker*, June 11, 2007.

Available online at: [http://www.newyorker.com/reporting/2007/06/11/070611fa\\_fact\\_max?currentPage=all](http://www.newyorker.com/reporting/2007/06/11/070611fa_fact_max?currentPage=all)

Charles Bukowski illustration that accompanies a letter to John Fante. In financial terms, an original document holds far greater value than a facsimile, even if the difference between the written content of the documents is negligible. As anyone who has worked with autograph collections will know, archival value and financial value are determined separately. A document can have both, neither, or one and not the other.

In contrast to analog literary archives, which have this inherent artifactual value, digital files do not; there is no digital original. Instead, any concept of originality of digital materials held in archives needs to be created and maintained through emerging standards for digital preservation and access. Though more of an art than a science, our professional community has settled on a process for agreeing upon financial value for analog papers but not yet an author's digital files. Which brings us to Susan Sontag.

Susan Sontag needs no introduction. Cultural critic, feminist icon, novelist....Sontag was a major literary and cultural figure and the definition of a public intellectual during the second half of the 20<sup>th</sup> century. In 2001, UCLA Library came to an agreement with Sontag to purchase her papers and massive book collection. Victoria Steele, then director of the Department of Special Collections, arranged for a donor to provide funding for the collection, which had been appraised for a high value. Containing Sontag's correspondence, writings, journals, photographs, and other materials, the papers represent a signature collection. Sontag died in 2004 and the following year we exercised our right of first refusal to purchase a significant addition to the papers, through her estate. The collection was arranged and described at this point and has been used heavily since.

In the winter of 2012, Sontag's literary agent contacted me on behalf of the estate, offering another six boxes of material documenting Sontag's early life and career that had been found. On a trip to New York to review the papers, I asked Sontag's son if his mother had used a computer. He told me that indeed she had and in fact he had two hard drives filled with the data from the computers she used during the last decade-plus of her life. I expressed a great deal of enthusiasm about the prospect of acquiring the data on the hard drives in addition to the paper-based materials initially under consideration and when I wrote to the Estate to finalize the acquisition, I requested they include the contents of the hard drives as part of the deal. At that time, they were unwilling to do so, wanting to review the digital holdings before agreeing to transfer them to UCLA, so the deal for the purchase of the paper materials went ahead as initially planned.

A few months later, I was contacted again by the agent to see if we wanted to purchase the hard drives in a separate deal. They arranged to have the contents of the hard drives copied and sent to us for our review. The timing was fortuitous, in that we were just establishing a digital archives program. My colleague Gloria Gonzalez had set up a forensics workstation, which she used to create a complete image of each hard drive. A group of us were then able to analyze the records for their archival value. What we discovered was truly exciting for anyone interested in Sontag and her milieu: nearly 18,000 email messages, correspondence, drafts and unpublished writings, photographs, and other documents taken from three computers she used; the digital materials provide a better view into the last decade of her life than the paper archive.

The first time I worked with Gloria to view the files, we pulled up an innocuously titled file that turned out to be a letter from Susan Sontag to the cultural critic Edward Said, trying to explain her decision to accept the Jerusalem Prize in 2001, about which Said had been a critic. The letter was written using Microsoft Word, in Times New Roman, on a Mac, which we were using to view the materials; I was immediately overcome by the rush of emotion so familiar to anyone who works with archival documents. This was surprising: I had always assumed the ways that computers tend to flatten expression would reduce or eliminate the emotional reaction we have to digital records, as compared to, say, a handwritten manuscript or letter. But here I was having those same reactions to what was at minimum a copy of a copy of a file that was once on a computer once owned by Susan Sontag, perhaps imagining that what I was seeing on the screen in front of me was, in fact, a close replica of how that file would have appeared to her as she was creating it.

I want to avoid a digression into the theories of Walter Benjamin to consider whether digital files can somehow be imbued with an aura in Walter Benjamin's sense of that concept but need to ask: did this digital file have an unexpected quality to it derived from an ability to tie it directly back to its creator? I believe what's really going on here is the thrill of discovery and privileged access that comes part and parcel when working with archival materials. But I raise the issue in general because it has some direct bearing to the main thrust of my paper here, a consideration of the value of digital archives. Clearly this letter and many others like it provide insight into Sontag, her politics, and relationships with fellow public intellectuals. But can a copy of a copy of a document have other values as well?

Turning directly to that question now, we decided for obvious reasons that we wanted to officially acquire the hard drives, but I was hesitant to buy the content for several reasons, as financial value for born digital archives remains a very unsettled business. As mentioned above, literary archives of an analog nature derive their value not only from their content and research value but largely from their artifactual—one is rarely, if ever, offered for sale the photocopies of a writer's correspondence even if the same content is being made available as the original. The whole concept of the forgery strikes right to the heart of the matter: what makes something valuable financially is its originality and its scarcity—the forger tries to mimic these elements to give something value that otherwise wouldn't have it. For archives and libraries, we want to acquire collections of one-of-a-kind items that no one else has. Yes, this helps us to support research, teaching, and understanding of the past, but we should also admit that we like being associated with the legacies of great writers and thinkers.

So what happens when that concept of originality gets taken off the table? One could argue that Sontag's Mac laptop would be singular—like Raymond Chandler's typewriter in UCLA's collection it would say something about her production environment and would provide a direct link back to the person herself. But in this case we are talking about data that has been copied at least twice. Furthermore the nature of digital technology allows for the EXACT copying of these files an infinite number of times; when copied correctly, there is no discernible or important difference between the ones and zeroes that sat on Sontag's Mac and the imaged

version now stored and backed up repeatedly on servers. In fact one pillar of our digital preservation strategy is Lots Of Copies Keep Stuff Safe, the LOCKSS framework. How can something devoid of originality be valuable in the same way as the physical materials commonly purchased by libraries and archives?

Our right of first refusal required the impartial analysis of an appraiser, who would use comparable values as the basis for their work. In this case, I was pretty sure we are breaking new ground and could not imagine either the estate or the library trusting the value an appraiser would place on digital archives, since no comparisons existed at the time or today as far as I know. Finally, we had a pretty good idea of the costs associated with acquiring, preserving, and making available for research most analog holdings. This too, is unsettled in the digital world, but at minimum we know it to be an expensive endeavor.

These were the bases of what I thought to be one of the better pieces of argumentative writing I have ever created, which I sent to the agent, urging the estate to donate the digital material, especially in light of the decade-long partnership UCLA had formed with them to preserve Sontag's legacy, which resulted in a substantial sums transferred to the estate. I proposed that estate could donate a copy of the Sontag digital files to UCLA, where it would be preserved and made available to those studying Sontag's life and work. We would ask for no restrictions to be put on the copies of the files still in the hands of the estate. The idealist in me struggles with the high values we place on literary archives in general and wants to live in a future where we can take consortial approaches to digital archives where we transcend institutional propriety and share the responsibilities for preservation and access.

The agent's response was swift and short. It intimated that because I hadn't made a financial offer for the digital files, we clearly didn't value them and he asked us to send the hard drives back with all due haste. Why did he refuse to define value only in financial terms? What I gathered from background reading is that some agents, and this agent in particular, seek out authors whose work he believes in. Eschewing bestsellers, he instead sees the work of his authors as continuing to pay off in the long term and to selling long after their death. Also, it's a clear and perhaps obvious point that authors and agents make their living by selling their writing; extending this logic to archives, digital or not, and filled with writings in many forms, makes reasonable sense. And in this case, the estate seemed unlikely to budge, leaving us with the unpleasant scenario of ceding our right-of-first-refusal and, more importantly, losing the opportunity to acquire what we considered an extremely important source for the Sontag archive.

Having reached an impasse, I conferred with our then-University Librarian Gary Strong and together we decided on a course of action about which I remain ambivalent: we cut a deal. Rather than requiring the estate to secure an appraisal, we offered a modest amount for the files, certainly relative to what we had paid for the paper archives. The estate accepted the offer and we got to work on a specialized agreement, with these pertinent details. In addition to the copies of the data already supplied by the estate, UCLA acquired the hard drives in the possession of the estate but allowed Sontag's son to retain them until his death, after which

they physically transfer to us. He can use them for his own purposes but cannot provide access to them or copy them for others. The agreement allows us to make all contents of the hard drives available to our users without restriction, according to the policies and procedures we use in our reading room. Like with the rest of the Sontag papers, intellectual property rights remain with the estate. Finally, a non-disclosure agreement prevents either side from revealing the amount of the transaction. I realize this might constitute a tease in the context of this paper, but the NDA was important from the estate's perspective and from my end, too, because I did not want these specific circumstances to be interpreted as setting a market for the digital component of future archives.

The benefits of this decision have been many and I will conclude with them as a postscript momentarily. My regrets stem from taking a baby-step down a problematic path for libraries and archives involved in preserving the record of cultural production in the digital realm. The values placed on literary archives have created a circumstance where only a small number of elite institutions can manage to build collections in this area. Even minor authors now expect big paydays for their archives, especially since they cannot get a tax break for self-generated materials. The toothpaste is out of the tube for analog archives and as we move further and further into the digital realm, I suspect the value of physical materials associated with writers will grow as it becomes scarcer. For digital archives, we have an opportunity to work differently and we should not move into this new era assuming an email has the same financial value as an Autograph Letter Signed; a Word file with tracked changes the same value as a heavily annotated and revised typescript. While they certainly share archival value, financial value does not necessarily follow and we need a new model that takes into account the lack of originality, ability for exact duplication, and complexity surrounding establishing and maintaining authenticity that fits the qualities and characteristics of the new media. If we can manage to change our paradigm, we have the opportunity to build a better way of connecting users to these materials. Imagine a future where institutions build collaborative systems for preservation, discovery, and access to archival holdings in a digital form. A great benefit to digital texts lies in their non-physicality—technology can liberate the user from needing to physically visit an archive. In addition, new types of analyses can be done on large groups of texts—the field of digital humanities has only begun to invent new ways to bring computational methods to help answer new types of research questions of our archives.

A proprietary approach to archives made sense in the physical world, but we have the capacity to serve our researchers in much greater manner by thinking outside the bounds of our own institutions. Of course, when you pay significant sums for archival holdings, there's much less incentive to think and act this way. But as our collections progress down the inevitable road of becoming more digital, what if we managed to direct the resources put into acquisitions into developing new mechanisms to serve a greater calling of making sure our holdings get used in ways that help us better understand our past and enrich our appreciation and understanding of our best writers and thinkers?

A cautionary note to my idealism relates to the authors themselves and those who represent them. Many authors have expressed ambivalence about their archives going to an institution

for a variety of reasons related to privacy and artistic pride. The prospect of a big payday for selling their papers provides an obvious material incentive, often stronger than the value proposition we make when seeking donations. But the reality is that the same forces that have disrupted the music and entertainment industry and how artists get paid for their work may be at play in archives as well. My hope is that our culture and society will develop new ways to ensure our writers and artists get fairly compensated for their work rather than employing mechanisms in the digital world that impede exchange of ideas and information and new methods of analysis.

As a way of concluding, I would like to share some further benefits of having acquired the Sontag digital archive. Aside from their archival and values, these materials proved quite valuable in other ways to our program. A major impetus yet undiscussed for me to acquire the Sontag digital files is that we have used them as a justification for and motivation to build our digital archives program. We have articulated a strategic direction to enhance our digital services and operations in order to meet the challenges posed by born digital materials and to maximize new technologies so they can help us improve our work to foster discovery of and access to our holdings. Having made a financial investment to acquire the Sontag files has made it easier to argue for resources to expand our program. We received approval for a digital archivist position on a project basis and the work Gloria and others have achieved over the past year has led to the approval of a permanent version of this position we will be advertising shortly.

Having purchased the digital materials using a strong agreement, we aggressively pursued means to provide access to it and our work quickly contributed to scholarship on Sontag. Most notably, we hosted Benjamin Moser, Sontag's authorized biographer, for an extended research visit. Ben spent an extensive amount of time reading and analyzing the Sontag email and other files, which had a significant impact on his research for the biography. Ben wrote a piece for the New Yorker blog that captured his reaction to doing research with this new type of archival record:

"reading papers and manuscripts is one thing. Looking through someone's e-mail is quite another, and the feeling of creepiness and voyeurism that overcame me as I sat with Gonzalez struggled with the unstoppable curiosity that I feel about Sontag's life. To read someone's e-mail is to see her thinking and talking in real time. If most e-mails are not interesting ... others reveal unexpected qualities that are delightful to discover. (Who would have suspected, for example, that Sontag sent e-mails with the subject heading "Whassup?") One sees Sontag, who had so many friends, elated to be in such easy touch with them one sees the insatiably lonely writer reaching out to people she hardly knew and inviting them to pay a call."

Even as he expresses some ambivalence, Moser's work has begun to articulate new insight that digital records provide into the thinking, lives, and creative process of their creators. Email has the seemingly conflicting qualities of being an incredibly intimate communication medium that expresses itself with relative visual uniformity. Several other researchers have utilized the files

as well, reminding us of why we do the work to capture and preserve archives in the first place—so that they can be used.

As I conclude this narrative, I am conscious our experience with the Sontag archive raises as many questions as it answers. Although I am very proud of the technical work we have completed on this material and the use it's been put to, my main impetus for reporting on this work has been to raise a set of questions we need to grapple with together as a profession and I would welcome pushback on the approach we took at UCLA if people have criticism that can help us move our theory and practice forward. One thing remains clear to me: we need to continue to evolve our practices and test our assumptions as our collections, programs, and approaches move further and further into a digital world.