How to Approach Hard Drives as Cultural Heritage

Johan Jarlbrink

Manuscripts, letters, and diaries are well established as documents in traditional archives. In collections of "papers" we can follow the lives and works of authors, scientists, politicians, and many lesser-known subjects. The documents have been donated by those who produced them, sometimes by relatives or the organizations they worked for. Many collections can be accessed by the public, or at least by researchers. Others are restricted in some way or another, often released after a preset date. Historians would have little to say about everyday life in the past if it were not for the files collected by archives.

We do not write very many letters on paper anymore. Susan Sontag's correspondence is a revealing example. She wrote letters and kept journals for most of her life, until she passed away in 2004. The materiality of her personal archive changed over time, however. In the mid-1990s she bought a PowerBook 5300, and later on a PowerMac G4 and an iBook. The library at UCLA did not know about her hard drives when they first acquired her personal papers in 2005, but in 2012 they got access to the three hard drives, with 18,000 emails, drafts, photographs, and other documents. The data is now part of Sontag's personal papers. Access is restricted, but not prohibited. Benjamin Moser, who published a biography about Susan Sontag in 2019, was able to go through some of the files as part of his research:

How to cite this book chapter:
Jarlbrink, Johan. "How to Approach Hard Drives as Cultural Heritage." In *Digital Human Sciences: New Objects—New Approaches*, edited by Sonya Petersson, 229–251. Stockholm: Stockholm University Press, 2021. DOI: https://doi.org/10.16993/bbk.j.

> [R]eading 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 [the librarian] Gonzalez [and] 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. […] One sees Sontag, who had so many friends, elated to be in such easy touch with them ("I'm catching the e-mail fever!"); one sees the insatiably lonely writer reaching out to people she hardly knew and inviting them to pay a call.[1]

Personal papers are even more personal when they are digital, it seems. And Moser did not even look at all the other kinds of data stored on hard drives—the web browser history, the words she googled, stored geolocations, metadata. Such information would be a gold mine for intellectual historians, but valuable also for media scholars researching everyday media life in the digital age. Scholars in other disciplines might want to investigate the software someone has used, or the file formats, viruses, the devices once connected to the computer—the research potential is enormous. These kinds of data are highly sensitive, of course, but so are many of the paper documents kept by archives. Why, then, is it so easy to access and read a private diary from 1897 but so difficult to find a web browser history from 1997?

The use of computers has been widespread since the 1980s, but hard drives as archives are still new to most memory institutions. Tom Hyry, the former director of the special collections at UCLA, explains that the incorporation of Sontag's hard drives into the library "raised technical, ethical, philosophical, financial, and practical issues that still seem new to the archival endeavor."[2] The preservation, curation, and presentation of someone's hard drive require technical solutions, skills, guidelines, and routines. One reason why the process has been slow for many institutions is perhaps that few scholars within the humanities and social sciences

---

[1] Benjamin Moser, "In the Sontag Archive," *The New Yorker*, January 30, 2014, https://www.newyorker.com/books/page-turner/in-the-sontag-archives.

[2] Tom Hyry, "On Digital Archives: Lessons from the Susan Sontag Hard Drives." Paper presented at the Society of American Archivists meeting, Cleveland, OH, August 20, 2015, http://nrs.harvard.edu/urn-3:HUL.InstRepos:40918991.

have taken an interest in these kinds of born-digital archives. This is unfortunate since hard drives are important records of everyday life, with a great potential as empirical sources.

The aim of this chapter is to exemplify what an investigation of a hard drive implicates, the methods needed to conduct it, and what kind of results we can get out of it. To focus the investigation, I will approach hard drives as records of everyday media use. I am less interested in the content of private emails and documents, what the photographs show or what secrets a web browser might reveal. I am interested in more general patterns of media use and how they change over time. To develop a media archaeological approach to the digital traces of everyday life (see the introduction to this volume), I will suggest a computer forensic method used as a media ethnographic tool. Computer forensics is a method developed to examine digital traces in order to establish a user's activities. Media ethnography is a field of inquiry researching media production and audiences in natural settings based on interviews, observations, media diaries, and field notes. The two methodological traditions are different but overlap. Computer forensics and media ethnography both take an interest in people's routines, the way they organize things, what they do, and how they do it. The approach I will demonstrate is a forensic investigation guided by media ethnographic themes, findings, and questions. Drawing from ethnographical studies of computers in everyday life, I will suggest the broad categories of time and space as fruitful starting points in digital excavations.

Hard Drives as Cultural Heritage

Media scholar Pelle Snickars wrote in 2010 that the hard drive was "our most central tool" and "the material base of our digital memory culture."[3] Yet, it is almost invisible in our daily lives and often buried behind several layers of plastic and circuit boards. Few scholars within the humanities or social science have paid much attention to it. Matthew Kirschenbaum is one exception,

[3] Pelle Snickars, "Hårddisken och samtiden [The Hard Drive Today]," in *The Story of Storage I*, ed. Lars Björk, Jānis Krēsliņš, and Matts Lindström (Stockholm: Kungliga biblioteket, 2010), 44 (my translation).

excavating the different layers of data from the perspective of digital literature. More on his pioneering work in the next section.

Snickars included servers in his discussion on hard drives in 2010. Ten years later it is perhaps these storage units, rather than personal hard drives, that form the material base of our memory culture. Documents, photographs, and messages are uploaded, sent, and accessed on online platforms. Few people save their Facebook posts on their own computer or phone. Most of us consume music and movies through streaming, and many documents are written and stored directly on platforms such as Google Drive. The biographies of future Susan Sontags cannot rely on personal hard drives alone. These outsourced storage units, however, are often beyond the reach of archives and libraries.

Personal computers as the prime storage units for digital information might be disappearing, but they are still important to memory institutions collecting the digital traces of the 1990s and 2000s. These were the decades when computers were domesticated, when they (at least in the West) became part of many people's everyday lives. The British Library and the US Council on Library and Information Resources raised the questions on how to collect and preserve personal hard drives in the late 2000s. In their reports they presented guidelines to support institutions and archivist in their work. Digital forensics was suggested as the prime method to capture and transfer data.[4] The decade following the two reports has seen many initiatives in the field. Museums, archives, and libraries have slowly started to incorporate data from hard drives into their collections. Computer forensics is nowadays "a standard practice in memory institutions for the preservation of digital storage media and born-digital records."[5] Yet, very few scholars within the humanities and social science have actually

[4] Jeremy Leighton John, Ian Rowlands, Peter Williams, and Katrina Dean, *Digital Lives: Personal Digital Archives for the 21st Century. An Initial Synthesis* (London: British Library 2010); Matthew G. Kirschenbaum, Richard Ovenden, and Gabriela Redwine, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* (Washington, DC: Council on Library and Information Resources, 2010).

[5] Thorsten Ries and Gábor Palkó, "Born-Digital Archives," *International Journal of Digital Humanities* 1, no. 1 (April 2019): 4, https://doi.org/10.1007/s42803-019-00011-x.

used and analyzed data from hard drives as primary sources in their research. Ries and Palkó (2019) write that there is a gap between memory institutions and archival science researchers on the one hand and scholars from the humanities and the social sciences on the other. There is a need to:

> enable GLAM institutions, institutional networks and infrastructures to develop their born-digital collections in meaningful ways, improve preservation formats, curation workflows, repositories, services, and access for researchers. This can only be achieved by cross-sectoral and interdisciplinary collaboration to support active research on born-digital collections.[6]

Hard drives from public figures such as Susan Sontag and Salman Rushdie are often key examples when preservation and access is discussed (for another fascinating case, see Amanda Wasielewski in this volume, about the excavation of the musical *RENT* on floppy disks left by the creator Jonathan Larson). If the strategies for collecting and preserving are based on data from intellectuals and writers, there is actually a risk that a textual bias will be built into the infrastructure. The archival strategies implemented today will have an impact on future research possibilities.[7] What I want to highlight here is that there are other—and overlooked—kinds of data to collect and explore.

Computer Forensics as a Method

Kirschenbaum has argued for a media specific reading of electronic literature. The digital format is an important part of what constitutes digital text. It is produced differently compared to the typewritten or printed text, it behaves differently, and it can be interpreted in different ways. Digital text is more than text on a screen.[8] My concern here is not digital poetry, but a basic understanding of digital storage is necessary also for an investigation of the hard

---

[6] Ries and Palkó, "Born-Digital Archives," 4.

[7] Agiatis Benardou et al., eds., "Introduction: A Critique of Digital Practices and Research Infrastructures," in *Cultural Heritage Infrastructures in Digital Humanities* (Abingdon: Routledge, 2018), 4.

[8] Matthew G. Kirschenbaum, *Mechanisms: New Media and Forensic Imagination* (Cambridge, London: MIT Press, 2008), xii–xiv.

drive as an archive. Data in digital form appears both fragile and stable. A sudden crash or a document not properly saved means that the data is lost for many users. A forensic investigation, however, will most likely be able to recreate such data.[9] To understand why, we need to know how data is stored on hard drives and how hard drives are managed within forensic investigations.

One way to describe the different appearances of digital objects is to differentiate between three basic forms or layers of the objects: the physical objects (inscriptions on a medium), logical objects (inscriptions read by software), and conceptual objects (as they appear on the screen).[10] Scholars within the humanities and social sciences have mostly dealt with digital objects in the third sense. Computer forensics, however, investigates all three forms. A basic definition states that it is concerned with "the examination of digital storage and digital environments in order to determine what has happened."[11] Tools and methods are developed in order to recreate (or monitor) events and actions based on digital traces. Some traces do not appear on the screen but can be identified as physical or logical objects. That is why all three layers or forms are important.

Most users know that the commands "delete file" or "empty trash" will remove a file's entry or address from the catalog but not delete the file itself from the hard drive. The space it takes up on the disk is flagged as available, but the data is only erased once it is overwritten with new data. With ever-increasing storage capabilities of disks, a space flagged as available might not be overwritten immediately. If the deleted file is overwritten, there are often copies in the form of temporary files or older versions saved elsewhere. Such files are created when files are modified, printed, copied, sent as attached files, *et cetera*.[12]

Since digital storage media are divided into clusters of a fixed length (often 4,096 bytes per allocated unit), individual files larger than a single cluster are stored in multiple places on the disk.

[9] Kirschenbaum, *Mechanisms*, chap. 2.
[10] Kirschenbaum, *Mechanisms*, 3.
[11] Joakim Kävrestad, *Fundamentals of Digital Forensics: Theory, Methods and Real-Life Applications* (Cham: Springer, 2018), 3.
[12] Kirschenbaum, *Mechanisms*, 52.

When a file is deleted, a part of it might be overwritten while other parts remain intact. These fragments can prove that a file existed even though most of the data is erased.[13]

In order to capture every fragment and every bit unaltered, the standard method within computer forensics is to copy the storage medium in a so-called bitstream. Similar to traditional archival practices, developed to keep documents and files in the same order as they were once arranged by those who produced them, a bitstream transfers every bit recorded in a linear sequence. To simply copy the files from one disk to another would miss the fragments and deleted files that do not show up among the indexed files in the graphical interface, and would add new metadata to the files. The computer or device should not even be turned on, as new data will be added as soon as the operating system starts running. Instead, the hard drive is taken out of the computer and is connected to a docking station. A bitstream captures every bit recorded on the disk and keeps the metadata intact. Since digital forensics was originally developed to support criminal investigations, it treats digital data like fingerprints and DNA on a crime scene, as evidence that should never be altered. The bitstream method makes the copy a stand-in for the original.[14]

Software for computer forensics usually lists the folders and files as they were arranged by the user, along with deleted files not yet overwritten, carved files (file fragments reassembled based on signatures in the code), and fragments in unallocated sectors of the disk. A USB stick might reveal a few hundred files, while a recent hard drive may contain tens of thousands of them. Most tools extract and highlight data of special interest to facilitate overview and orientation. Autopsy (4.11.0), the tool I am most familiar with, generates lists of image files, video and documents, EXIF metadata (the cameras used to take photos, among other things), encrypted files, accounts, emails and addresses, and very large files. A search function makes it possible to locate files and fragments containing particular keywords. The tool is obviously designed with criminal investigators in mind, for those looking

[13] Kirschenbaum, *Mechanisms*, 52.
[14] Kirschenbaum, *Mechanisms*, 53.

for specific files, contacts, or addresses. An example from *Digital Evidence and Computer Crime* (2011) is typical:

> <A HREF="http://www.google.com/search?hl=en&lr=&ie=ISO-8859-1&q=human+poison+herbs" ADD_DATE="1049641841" LAST_VISIT="1049642467" VISITATION_COUNT="3 "OBJECT_TYPE="LINK">[15]

Scholars interested in general patterns of user behavior and how they change over time have more limited options. The timeline view in Autopsy (Figure 1) gives a useful overview, though, showing events (files created, accessed, modified, deleted, web activity, installed programs, *et cetera*) in a sequence, with the possibility to zoom in (on events at a specific date, minute, second) or out (indicating changes from year to year).

An even better option might be to export the metadata to Excel, where it can be filtered and grouped depending on the research interest. The metadata for an individual file usually includes file name, extension (file format), when it was created, modified, accessed, and deleted (if it was); its size, address, hash value (an id derived from the data, making it possible to locate duplicates and similarities), and its full path. Based on this metadata, researchers can examine users' information management, how files are organized into folders, when different kinds of software were installed, how the generation of different file formats change over time, when files in specific folders were created and modified, and so on.

Files of special interest may also include log files, browser history, playlists, and traces of various devices connected to the computer, such as printers, scanners, USB sticks, and phones. Forensics is based on the idea that "[e]very contact leaves a trace," and this is no less true for computer forensics.[16] This does not mean that every contact is traceable, however. What data and metadata is available depends on the operating system and the file system. Older generations of Apple computers, for example, did not add extensions to the file name or timestamps for "last accessed."

---

[15] Eoghan Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet* (Burlington: Academic Press, 2011), 599.

[16] Kirschenbaum, *Mechanisms*, 49.

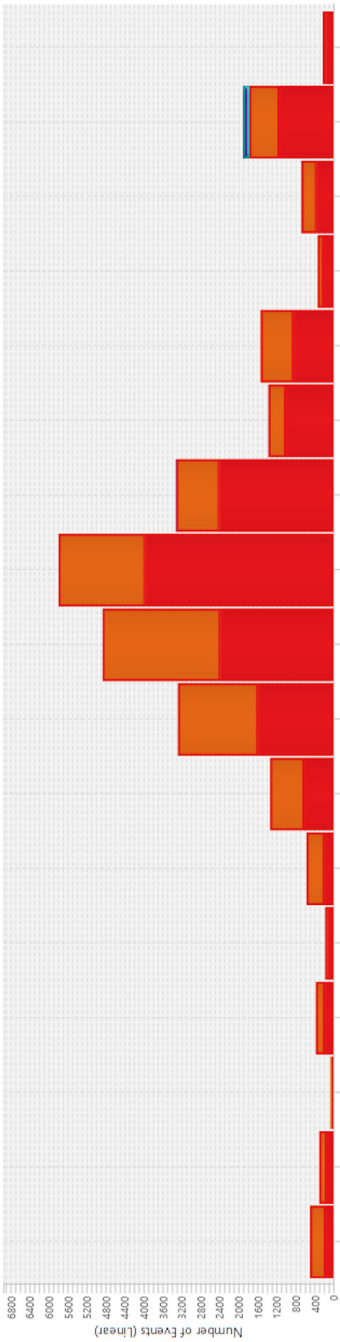

Figure 1: Timeline created with Autopsy 4.11.0, based on an Apple iMac G3 (id TEKS0047646) in the collections of the National Museum of Science and Technology, Stockholm, Sweden. Orange represents files created, red files modified, and blue miscellaneous files, in this case photographs. Copyright: Author. License: CC BY 4.0.

Deleted and carved files may lack most of the metadata.[17] Still, many aspects of a person's digital life can be recreated based on the files, logs, and metadata. Part of it is very personal; other parts may be less sensitive. The rise and fall of MP3 files on personal computers, or the history of preferred search engines, or the cycles of constantly new software updates, or the domestic ecosystems of connected devices—these issues might be examined without sensitive information being required or revealed.

Mapping Time and Space

In order to sketch potential research topics, where forensic methods and findings can complement and stimulate media ethnographical approaches, I will mainly draw on two previous studies of computers and digital communication in everyday life, Elaine Lally's *At Home with Computers* from 2002, and Maria Bakardjieva's *Internet Society: The Internet in Everyday Life* from 2005. They were not written as historical accounts of media use—but now that is what they are. Lally interviewed 95 individuals in 31 households in Australia. Most of the interviews covered experiences from the second half of the 1990s. Bakardjieva interviewed 23 respondents living in Canada about media use in the late 1990s and early 2000s. Lally took an interest in people's computer interaction in a broad sense; Bakardjieva was primarily focused on respondents' experiences of the internet, but she also described domestic computer usage in general.

Most of the respondents interviewed by Lally and Bakardjieva lived in families with only one computer. This computer was often their first. For many of them, the computer represented a "shared space," a technology available for every family member, but (mostly) for one person at a time. A shared computer called for more or less developed rules of "time-zoning."[18] The computer was used by certain family members for certain things at certain times, depending on internal hierarchies of power, values, skills, and needs. Thus, much of what Lally and Bakardjieva described

[17] Casey, *Digital Evidence,* 588–591.

[18] Maria Bakardjieva, *Internet Society: The Internet in Everyday Life* (London: Sage, 2005), 151.

was how the use of computers was organized and restricted in time and space.[19]

A hard drive itself is also a shared space, filled with data in temporal layers. A forensic analysis reveals how this space is divided and used, when, for what, and by whom. Lally and Bakardjieva were less interested in the content of people's emails, exactly what the respondents searched for online, or what kind of work or entertainment they were engaged in. What they examined was the way computer interaction was organized socially. My intention here is to build on their work in order to map potential overlaps between forensic inquiries and ethnographic approaches. What kind of questions derived from the ethnographic analysis can we investigate further by forensic means? What kind of forensic findings might be contextualized in ethnographical studies?

Time

Computers that end up in museums or archives have unique histories as objects in specific social settings. One way to approach a computer as a historical artifact is to examine the "cultural biography" of the object and its journey in time and space. When anthropologist Igor Kopytoff launched the idea of cultural biographies of things in 1986, he stated that we can ask similar questions about things as we do about people:

> Where does the thing come from and who made it? What has been its career so far, and what do people consider to be an ideal career for such things? What are the recognized "ages" or periods in the thing's "life," and what are the cultural markers for them? How does the thing's use change with its age, and what happens to it when it reaches the end of its usefulness?[20]

What is specific about a computer is that it keeps track of its own history. We can ask the former users about its history, and we

[19] Both of them have chapters on "Temporal Rhythms of the Computerized Home" (Lally) and "Making Room for the Internet" (Bakardjieva).

[20] Igor Kopytoff, "The Cultural Biography of Things: Commoditization as Process," in *The Social Life of Things: Commodities in Cultural Perspective*, ed. Arjun Appadurai (Cambridge: Cambridge University Press, 1986), 66–67.

can trace its history from within the object itself. In a biographical investigation of a computer it would be fruitful to distinguish different "ages," what characterized these, and how and why a computer transitioned from one age to the next. Jonathan Stern has pointed out that the aging of computers is in most cases a process driven by repeated introductions of new software. A computer bought a few years ago might be running perfectly fine with its original software, but it is not powerful enough to run new software. Thus, the computer *becomes* old when new software is introduced.[21] The history of installed and updated software can easily be tracked in log files and in the metadata of program files. Such metadata can also show the old within the new. The timeline in Figure 1 is based on a hard drive manufactured in 1998 and bought in 1999—and reveals that many of the files in the preinstalled software packages were created in the early 1990s. The users might think that what they buy is brand new, but much of it is old stuff packaged in a new box.

Several users interviewed by Lally and Bakardjieva were well aware that new media ages fast: "I bought a 486—at that time it was the best and today it is already old and out of date."[22] "[I]t seemed to be out-of-date as soon as we got it. It probably wasn't really quickly, but it just seemed really quick."[23] In another case, it was a new computer at work that made the home computer from the previous year look ancient.[24] Some of the old, slow, and discarded computers were given a second chance, however. Computers no longer used by adults were sometimes given to children, and some men handed over their old machines to their wives. One of the most common ways to acquire a home computer was to purchase (or just take over) an old computer from work, colleagues, and friends.[25] Thus, early adopters spread the technology to laggards in their social networks: "Non-professionals and

[21] Jonathan Sterne, "Out with the Trash: On the Future of New Media," in *Residual Media,* ed. Charles R. Acland (Minneapolis, MN: University of Minnesota Press, 2007), 22–23.
[22] Bakardjieva, *Internet Society*, 93.
[23] Elaine Lally, *At Home with Computers* (Oxford: Berg, 2002), 85.
[24] Lally, *At Home*, 117.
[25] Lally, *At Home*, 74–88.

'poor cousins' take up the computer waste and put it to uses of their own."[26]

A standard life journey of a computer takes the following route in Sterne's account: "it travels through categories from new, to useful, to obsolete, to unused, to trash."[27] As exemplified in the interviews, however, there might be several detours on the journey. A forensic analysis can trace the journey and reveal several temporal layers of data, from multiple users or owners. Such an analysis can show how usage changed over time. A new computer is most likely used differently from one that is four years old, and different owners use it for different things. The timeline in Figure 1 indicates some of the changes on a macro level, but a thorough analysis of the underlying data would provide us with a detailed biography as well as different user profiles. Much of the social context would be missing from the analysis, but it could be a useful addition to traditional ethnographical approaches.

New media studies are often occupied with the latest version, new gadgets, and new practice. A computer forensics approach may provide an alternative and reveal hardware and software in use long after new versions were launched. Historian of technology David Edgerton made the claim in *The Shock of the Old* (2006) that "many things we think of as old remained in practical use for longer than our future-oriented accounts of technological history allow."[28] What is true for spaceships and sewing machines is most likely true for computers: many of the technologies in use are surprisingly old. Some users transfer old software to the new computer when they upgrade. Others install emulators to be able to play old games on new machines. A forensic analysis of metadata makes it possible to distinguish the temporal layers.

We can also investigate a different kind of temporality, the daily rhythms of computer interaction. Such an interest was an important part of computer forensics right from the beginning. When astronomer Clifford Stoll in Berkeley tried to track a hacker in 1986, breaking into various military networks all over the USA

[26] Bakardjieva, *Internet Society*, 94.
[27] Sterne, "Out with the Trash," 23.
[28] David Edgerton, *The Shock of the Old: Technology and Global History Since 1900* (New York: Oxford University Press, 2006), 29.

via the computer in Stoll's own lab, he noticed that the hacker entered the systems at specific hours: "On the average, the hacker showed up at noon, Pacific time. Because of daylight savings time, I could stretch this to 12:30 or even 1 P.M., but there was no way that he was an evening person."[29] Most hackers worked evenings and late nights due to the lower costs of data traffic, but not this one. Could this indicate that the hacker came from overseas? Early afternoon in California was late night in Europe. Stoll broadened his search and managed to track him down—in Hannover, Germany. Stoll's pioneering work laid the ground for what is today computer forensics. His approach to time might be valuable also for scholars analyzing the mundane rhythms of digital life.

How the families studied by Lally and Bakardjieva used their computers varied with the seasons, with school semesters and breaks, weekdays, and weekends. How to divide, regulate, and spend computer time was one of the most frequently reoccurring issues. According to the domestic moral economy of time, some activities were encouraged, while others were restricted:

> this hierarchy may be institutionalized in explicit rules ("homework comes before games") and conventions ("the person whose homework deadline comes first has first turn"), but is also open to negotiation (adults can stay up later than children so the homework can sometimes come before adult income-generating work).[30]

Members of families with only one computer allocated time in the same way they allocated space. They had to define rules, and negotiate and divide time slots depending on authority and need. Many of the adults struggled to separate work and leisure time. Some spent evenings on work-related tasks in front of their computers, but computers from work were also used for activities not related to work.[31] How much computers were used, by whom and when, also changed over time, depending on the age of family members and the age of the computer itself. Initially, many families

[29] Clifford Stoll, *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage* (New York: Doubleday, 1989), 138.
[30] Lally, *At Home*, 131.
[31] Bakardjieva, *Internet Society*, 96.

used them in the same way as vacuum cleaners—turned on only when used. Over time they came to resemble fridges instead—never turned off.[32] Aged computers were sometimes used less and less, however, as they became slow and not compatible with new software.

How and when computers have been used has changed with the internet connections available. The tariffs were higher for daytime surfing well into the 2000s, explaining why some people chose to download heavy files, games, and large software during early mornings or late at night. When internet connections used the same landline as the regular phone, the time online had to be limited in order not to block phone calls.[33] New connections and flat rates have meant that the time spent online has steadily increased. Jonathan Crary has pointed out that modern media have rearranged our daily rhythm, how we spend our time, when we sleep, and how much (a related idea concerns the "acceleration of 'the pace of life'," see Jonas Stier in this volume). Computers in standby mode and always connected have turned weekends into workdays, while an endless supply of content online makes us stay up late at night.

> The notion of an apparatus in a state of low-power readiness remakes the larger sense of sleep into simply a deferred or diminished condition of operationality and access. It supersedes an off/on logic, so that nothing is ever fundamentally "off" and there is never an actual state of rest.[34]

A forensic analysis of hard drives could ground this analysis in actual user behavior, or reveal patterns that are more complex. What does a daily rhythm of computer use look like? How does it differ between individuals in different contexts? Does it change over time? How?

Figure 2 shows the number of files created at every hour in a personal folder. It is stripped of most of the details, but indicates a personal rhythm nevertheless. One possibility is to filter the data

---

[32] Lally, *At Home*, 127.

[33] Bakardjieva, *Internet Society*, 148, 151.

[34] Jonathan Crary, *24/7: Late Capitalism and the End of Sleep* (London, New York: Verso, 2013), 13.

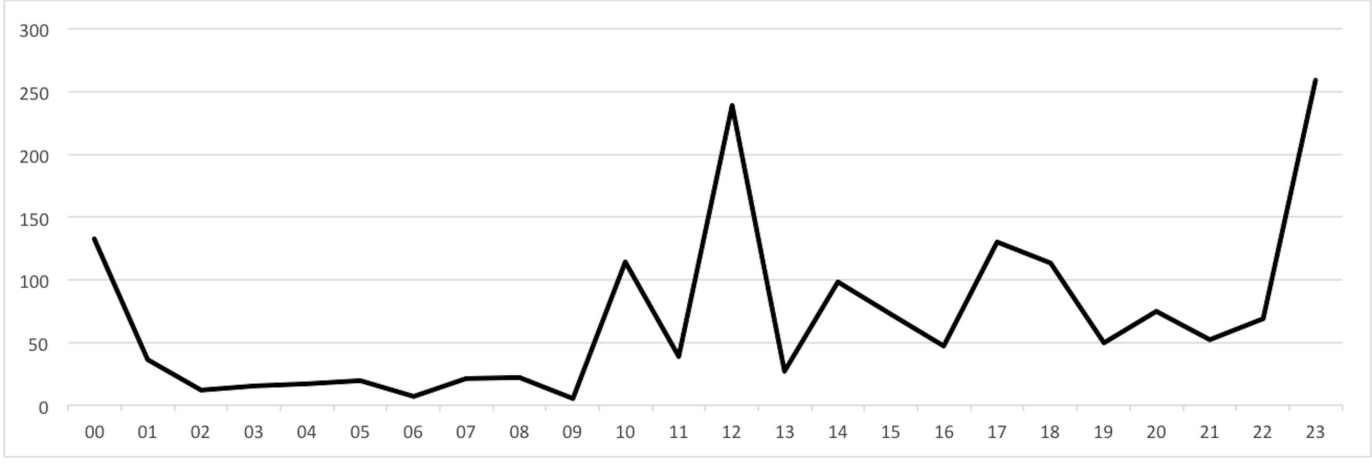

Figure 2: Timeline created in Excel, based on file metadata from a personal folder on an Apple iMac G3 (id TEKS0047646) in the collections of the National Museum of Science and Technology, Stockholm, Sweden. Copyright: Author. License: CC BY 4.0.

and examine timestamps for specific file formats. Do MP3 files follow a different rhythm than Word files and PowerPoints? We could also compare different personal folders within a family. If a folder is divided into "Private" and "Work," we can compare the timestamps in order to understand divisions and overlaps between work and leisure time. Data from the forensic examination can be explored further with ethnographical methods, and vice versa. What did you do around lunchtime? What kept you up late at night? What happened between 7 and 10 pm? Figure 2 gives a basic insight into what is possible, but there is a lot more that can be done.

Space

The life journey of a computer, as sketched by Jonathan Sterne (from new to trash), usually corresponds to movements in space: from work to home, from the living room or home office into the children's room, from bedroom to basement, and from visible areas into closets.[35] The geography of a family home is often based on notions of shared space and private, adult areas and underage, female and male.[36] Where a computer is placed has consequences for when it can be used, by whom, and for what. Parents might choose a specific location for the computer in order to control how children are using it. Some rooms are "neutral" and used by everyone, like living rooms and hallways, while others, such as bedrooms, are private. One of Bakardjieva's respondents explained that "the living room was where everybody could have access to it. Because we just have the one [computer], if we had more than one, it may have made sense to have it in a study, or a bedroom, but just the one, we put it, you know, central."[37]

In some families an individual member claims "ownership" of the computer and restricts other members' access by placing it in a private room. When such formal or informal ownership changes, the computer is often moved to another room.[38] How the computer is used may change with its physical location: private or noisy

[35] Sterne, "Out with the Trash," 25.
[36] Bakardjieva, *Internet Society*, 139.
[37] Bakardjieva, *Internet Society*, 151.
[38] Lally, *At Home*, 138.

stuff (games and music), or work that requires silence and concentration, in closed-off areas, and less private tasks in a shared space.[39]

The hard drive itself is a space that also needs to be shared and divided. In families with only one computer it is often divided into a common desktop and personal folders. To put personal files directly on the desktop is to colonize common space. One of Lally's respondents complained that her husband had "created a folder called 'Masters' for his university work which is outside all the individual family members' own folders." She saw this as "proprietorial": "You should have your 'Masters' inside 'Thomas'."[40] How the memory space available is used by the different family members can tell us something about formal or informal ownership. Take the personal folders of this 4 GB hard drive as an example (Table 1):

Who "owned" this computer? From file size alone, we can at least say that person 2 was the one who was taking up most storage space with his or her personal stuff. The files in the other persons' folders do not come close when it comes to size. If we add time to the equation, we would see that person 1 was actually creating most of the files in the personal folder in year one—but next to nothing the following years. Person 2 created most of the files in years two and three. Person 3 was creating files from year one until year seven—but small Word files that do not take up much space. "Ownership" and use usually change over time, and a forensic analysis makes it possible to follow changes in detail.

Many of the files on a hard drive are not personal at all. They come with the computer when it is bought, or are stored on the

Table 1: The size of personal folders, based on metadata from an Apple iMac G3 (id TEKS0047646) in the collections of the National Museum of Science and Technology, Stockholm, Sweden.

Person 1	169.2 MB
Person 2	1012.9 MB
Person 3	15.2 MB

[39] Bakardjieva, *Internet Society*, chap. 6.
[40] Lally, *At Home*, 138.

hard drive when new software is installed or old ones updated. The hard drive is a space that users share with software manufacturers. This part of the space is also worth an excavation.

The field of software studies (in Lev Manovich's version) is concerned with "the role of software in contemporary culture, and the cultural and social forces that are shaping the development of software itself."[41] Matthew Fuller writes about methods within the field that "software studies approaches might characteristically tend to identify specific algorithms, articulate their genealogy, recognize and work with their characteristics, and see them as part of a larger assemblage."[42] Forensic investigations of personal hard drives might provide software studies with individual stacks of software, what they were once used for, how they were used in combination, what was installed but immediately uninstalled, and so on. Another possibility is to analyze the way space taken up by software was once customized by individual users. To design and download "skins" for the MP3 player was very common around 2000, as were various icons used for ICQ accounts, *et cetera*.[43] Plug-ins and personal settings are other examples.

To map time and space in the ways I have suggested here is just meant to exemplify possible research. Other aspects to explore include what Lally describes as "the domestic ecology of objects," the way computers become connected to printers, floppy disks, CDs, USB sticks, modems, phones, and so on.[44] From here we can follow how domestic networks are connected to the world outside.

Concluding Remarks: Institutional and Methodological Challenges

Before we can follow any of the traces, they must be available for research. Unless researchers get their hands on hard drives

[41] Lev Manovich, *Software Takes Command* (New York, London: Bloomsbury, 2013), 10.

[42] Matthew Fuller, "Software Studies Methods," in *The Routledge Companion to Media Studies and Digital Humanities*, ed. Jentery Sayers (New York: Routledge, 2018), 251.

[43] Jeremy Wade Morris, *Selling Digital Music, Formatting Culture* (Oakland, CA: University of California Press, 2015), 55.

[44] Lally, *At Home*, chap. 9.

themselves, archives, libraries, and museums are logical repositories for this kind of material. These institutions have a long tradition of collecting personal papers and objects, and know the importance of protecting sensitive information. Analog and digital collections have much in common, but there are important differences. A home computer with an internet connection is likely to record data about everyday life in all its diversity, whether the user is aware of it or not. Data from and about other people will be recorded as well, without them knowing it. Personal papers are often more selective, at least those that reach archival institutions. Papers once shredded by a donor are gone forever, while many files deleted on computers can be recovered. Archives, libraries, and museums need to develop strategies, routines, and technical skills in order to handle born-digital collections in ways that are legal, are archivally sound, and protect private integrity. Donors must be informed about the possibility that hidden data might be recovered.[45]

Forensic procedures follow archival standards prescribing authenticity and provenance. Yet, the completeness that is a requirement in a criminal investigation can cause practical problems for archival institutions: a disk image contains more data than some of them can handle. Sensitive information, about the original user(s) as well as others, can be hidden among thousands of files. Some of it can be located automatically with forensic software, such as addresses, bank accounts, and social security numbers. To locate other kinds of data requires manual processing—and much time. In an ideal situation this is done in close contact with the donor, but this is not always possible. Some institutions choose to keep much of the unsorted and potentially sensitive data in a secure and restricted "dark archive," while providing access to a selection of nonproblematic files only—this is the way Salman Rushdie's hard drives are curated and made available.[46]

---

[45] Kirschenbaum, Ovenden, and Redwine, *Digital Forensics*, 29–39; Alyssa Hamer, "Ethics of Archival Practice: New Considerations in the Digital Age," *Archivaria* 85 (Spring 2018): 156–179.

[46] Ben Goldman and Timothy D. Pyatt, "Security Without Obscurity: Managing Personally Identifiable Information in Born-Digital Archives," *Library and Archival Security* 26, no. 1–2 (2013): 44–45, 50, https://doi.org/10.1080/01960075.2014.913966.

Others avoid the workload generated by forensic methods, and let donors themselves transfer the files they want to include in the digital repository via online platforms. This, however, may exclude and distort metadata and the context that makes it possible to determine how and when files were created, where they were created and stored, and so on.[47]

It makes sense for a library to primarily collect selected text files from well-known authors. Other institutions, such as museums of ethnography or technology, should consider capturing and preserving a broader range of data. As I have shown in this chapter, discarded hard drives have much to tell us about the way we live our lives, about daily routines and rhythms, and about how technology became part of everyday life. Important clues to explore can be found in all kinds of files, but also in the metadata. This category of data is often (but not always) less sensitive than the actual content of files. The file metadata from a single hard drive is not big data; it can easily be processed manually if parts of it need to be redacted.

If a disk image has already been prepared by archivists when the hard drive is donated, researchers do not need to repeat the process. Nevertheless, they need to think like a forensic investigator and need to be aware of the principles of digital storage and how different file systems generate different kinds of metadata. A disk image captures every trace on the storage medium, but not all traces ever left. There is no way of knowing what is overwritten and missing from the record. Still, researchers interested in issues related to routines, time and information management, the domestication of technology, and the history of new media have much to gain from a forensic investigation of a hard drive. An interview relies on an interviewee's self-understanding and what he or she can remember. A disc image represents a different kind of memory, a record of the microscopic details of everyday media use. Ethnography and computer forensics have different roots, but they can work in concert.

---

[47] Katinka Ahlbom (head of the department for manuscripts, maps, and images, National Library of Sweden), personal communication November 21, 2019.

Acknowledgments

The research for this chapter was conducted within the project *Digital Models: Techno-Historical Collections, Digital Humanities & Narratives of Industrialisation*, funded by the Royal Swedish Academy of Letters, History and Antiquities. My gratitude also goes to Jim Robertsson at Humlab, Umeå University, for guiding me through the technicalities of the forensic investigation.

References

Bakardjieva, Maria. *Internet Society: The Internet in Everyday Life*. London: Sage, 2005.

Benardou, Agiatis, Erik Champion, Costis Dallas, and Lorna Hughes, eds. "Introduction: A Critique of Digital Practices and Research Infrastructures." In *Cultural Heritage Infrastructures in Digital Humanities*, 1–14. Abingdon: Routledge, 2018.

Casey, Eoghan. *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet*. Burlington: Academic Press, 2011.

Crary, Jonathan. *24/7: Late Capitalism and the End of Sleep*. London, New York: Verso, 2013.

Edgerton, David. *The Shock of the Old: Technology and Global History Since 1900*. New York: Oxford University Press, 2006.

Fuller, Matthew. "Software Studies Methods." In *The Routledge Companion to Media Studies and Digital Humanities*, edited by Jentery Sayers, 250–257. New York: Routledge, 2018.

Goldman, Ben, and Timothy D. Pyatt. "Security Without Obscurity: Managing Personally Identifiable Information in Born-Digital Archives." *Library and Archival Security* 26, no. 1–2 (2013): 37–55. https://doi.org/10.1080/01960075.2014.913966.

Hamer, Alyssa. "Ethics of Archival Practice: New Considerations in the Digital Age." *Archivaria* 85 (Spring 2018): 156–179.

Hyry, Tom. "On Digital Archives: Lessons from the Susan Sontag Hard Drives." Paper presented at the Society of American Archivists meeting, Cleveland, OH, August 20, 2015. http://nrs.harvard.edu/urn-3:HUL.InstRepos:40918991.

Kirschenbaum, Matthew G. *Mechanisms: New Media and Forensic Imagination*. Cambridge, MA, London: MIT Press, 2008.

Kirschenbaum, Matthew G., Richard Ovenden, and Gabriela Redwine. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Washington, DC: Council on Library and Information Resources, 2010.

Kopytoff, Igor. "The Cultural Biography of Things: Commoditization as Process." In *The Social Life of Things: Commodities in Cultural Perspective*, edited by Arjun Appadurai, 64–91. Cambridge: Cambridge University Press, 1986.

Kävrestad, Joakim. *Fundamentals of Digital Forensics: Theory, Methods and Real-Life Applications*. Cham: Springer, 2018.

Lally, Elaine. *At Home with Computers*. Oxford: Berg, 2002.

Leighton John, Jeremy, Ian Rowlands, Peter Williams, and Katrina Dean. *Digital Lives: Personal Digital Archives for the 21st Century. An Initial Synthesis*. London: British Library 2010.

Manovich, Lev. *Software Takes Command*. New York, London: Bloomsbury, 2013.

Morris, Jeremy Wade. *Selling Digital Music, Formatting Culture*. Oakland, CA: University of California Press, 2015.

Moser, Benjamin. "In the Sontag Archive." *The New Yorker*, January 30, 2014. https://www.newyorker.com/books/page-turner/in-the-sontag-archives.

Ries, Thorsten, and Gábor Palkó. "Born-Digital Archives." *International Journal of Digital Humanities* 1, no. 1 (April 2019): 1–11. https://doi.org/10.1007/s42803-019-00011-x.

Snickars, Pelle. "Hårddisken och samtiden [The Hard Drive Today]." In *The Story of Storage I*, edited by Lars Björk, Jānis Krēsliņš, and Matts Lindström, 44–51. Stockholm: Kungliga biblioteket, 2010.

Sterne, Jonathan. "Out with the Trash: On the Future of New Media." In *Residual Media*, edited by Charles R. Acland, 16–31. Minneapolis, MN: University of Minnesota Press, 2007.

Stoll, Clifford. *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*. New York: Doubleday, 1989.