1

"Every Contact Leaves a Trace": Storage, Inscription, and Computer Forensics

> It should always be emphasized that physical facts are not less significant simply because the unaided eye cannot see them.
> —ALBERT S. OSBORN, *QUESTIONED DOCUMENTS* (SECOND EDITION), 1929

> Each diskette is a small (about 5-inch diameter) plastic disk coated so that information may be stored on and erased from its surface. The coating is similar to the magnetic coating on a recording tape. The diskette is permanently sealed in a square black plastic cover which protects it, helps keep it clean and allows it to spin freely. This package is never opened.
> —*THE DOS MANUAL*, APPLE COMPUTER INC., 1980

> Visibility itself is not a measure of inscription, modification of the substratum is.
> —MARCOS NOVAK, "TRANSTERRAFORM" (UNDATED, ONLINE)

The most uncompromising statement on the materiality of digital media I know is a Department of Defense document labeled DoD 5220.22-M, the Operating Manual for the National Industrial Security Program.[1] Initially

[1] Available at http://www.dtic.mil/whs/directives/corres/html/522022m.htm as well as many other locations online.

published in 1991, it seeks to establish and standardize security practices for handling classified information at the juncture between government and industry. Some eighty pages in we encounter the Clearing and Sanitization Matrix, a table listing numerous varieties of magnetic and optical storage media together with DoD-sanctioned methods for removing data stored on each type. The options range from simple overwrites (recording random or arbitrary bits on top of existing information) to various levels of degaussing (using magnetic fields to neutralize the polarity of the magnetic media, thereby sanitizing it), to Option M, available for all optical and magnetic media: "Destroy—Disintegrate, incinerate, pulverize, shred, or smelt." Some sense of what this means in actual practice may be conveyed by the following colorful account, posted to a USENET newsgroup:

> When I had to declassify disk drives in 1987, the NSA suggested that I take them out to the parking lot, and run them over with a tank. . . . I told him that the Pentagon parking lot had about 12,000 cars, but no tanks. His second choice was that we put the drive on top of a research magnet the Navy had. . . . I don't know what the field strength of that magnet was, but it had big warning signs all over the building. You had to take off everything metal just to go into the same room. The magnet consumed 186 volts at 13,100 amps. That's about 2.5 megawatts. We left it there for about a minute and a half. The field physically bent the platters on our 14-inch drive.[2]

The DoD's Clearing and Sanitization Matrix offers a bracing counterpoint to the first wave of academic writing on electronic textuality, with which it is exactly contemporary. While media scholars and literary theoreticians were feeling their way toward metaphors and neologisms designed to capture something of the fleeting quality of the flickering signifiers on their screens, bureaucrats at the DoD were wringing their hands over electronic data's troubling penchant for remanence—defined by an influential National Computer Security Center study as "the residual physical representation of data that has been in some way erased."[3] They were enumerating the relevant variables in a matrix while experimenting with a myriad of techniques designed to render

[2] Posted by David Hayes to the comp.periphs.scsi newsgroup, 24 Jul 91 05:07:01 GMT, as "Re: How many times erased does DoD want?"

[3] NCSC-TG-025, *A Guide to Understanding Data Remanence in Automated Information Systems*. Widely available online: http://crypto-systems.com/datarem.html.

discarded information invulnerable. Taken together, the academy and the DoD reveal two starkly different attitudes towards the textual condition of electronic objects circa 1991 (one year prior to the production of *Agrippa*) and ask us to develop an approach capable of accounting for the ways in which electronic data was simultaneously perceived as evanescent and ephemeral in some quarters, and remarkably, stubbornly, perniciously stable and persistent in others.

Stored Programs and Screen Essentialism

A document like the Clearing and Sanitization Matrix exists because of a particular tradition of computing: the stored program, which entails both physical and logical separation of the processing unit from a computer's memory, encompassing both data and instructions for operating on data—the literal stored program. John von Neumann's 1945 "Draft Report on the EDVAC" remains the single most influential and complete articulation of the stored program concept, even if it is not the sole progenitor. The Draft Report effectively dictates that there is no computation without data's representation in a corresponding physical substratum, the specifics of which very quickly get us into a messy world of matter and metal whose minute particulars seem conspicuously at odds with the equations and schematics dominating the rest of von Neumann's text: "[I]nstructions must be given in some form which the device can sense: Punched into a system of punchcards or on teletype tape, magnetically impressed on steel tape or wire, photographically impressed on motion picture film, wired into one or more fixed or exchangeable plugboards—this list being by no means necessarily complete."[4]

A year earlier John Mauchly, who worked with von Neumann at the Moore School on the ENIAC and who with Wallace Eckert would soon leave Penn to build the UNIVAC, had posited storing numeric data on "disks or drums which have at least their outer edge made of a magnetic alloy."[5] It is hard to think of such a scheme as "writing" in anything but the most generic sense—probably we are led to think in terms of materials science and fabrication

[4] "First Draft Report on the EDVAC," http://www.virtualtravelog.net/entries/2003-08-TheFirstDraft.pdf.

[5] Quoted in Paul E. Ceruzzi, *A History of Modern Computing* (Cambridge: MIT Press, 1998), 22.

instead. Yet, as scholars such as Frank Salomon have noted in their work on the khipu, the ancient Incan information recording device that stored data in knotted cords, the exclusive identification of writing with phonetic sign systems has been challenged in a number of quarters.[6] The magnetic storage devices envisioned by von Neumann, Mauchly, and others have become the preeminent information storage devices of our own day; indeed, so thoroughly integrated into our daily lives are magnetic recording media that we routinely embed them in an even older substrate, paper, so that a disposable printed card carelessly slipped into a pocket becomes the commodity token which admits me to my local transit system and debits the cost of my ride from a value recorded on the magnetic strip.

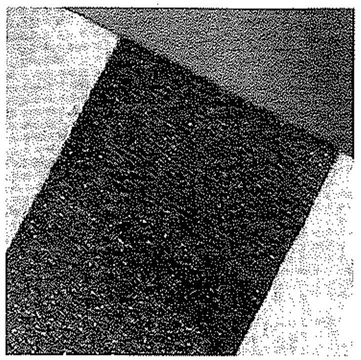


Figure 1.1 Magnified image of a small portion of the magnetic strip on a Washington, D.C. Metro fare card after the application of MagView developer's fluid. Striated data tracks are most clearly visible near the top edge, perpendicular to the length of the strip. Photograph by the author.

Though normally invisible to human eyes, the magnetic recording on such a card is indisputably an inscription, as is apparent after the application of aerosolized ferrite oxide, which makes the tracks and data patterns visible (see figure 1.1).

[6] Frank Salomon, *The Cord Keepers: Khipus and Cultural Life in a Peruvian Village* (Durham: Duke University Press, 2004), 23–30. The crucial intervention is Geoffrey Sampson's, who in his book *Writing Systems* (Stanford University Press, 1985), elaborates a concept of the graphic sign known as semasiography to sit alongside of the glottographic systems of which the alphabet, where signs stand in for phonetic speech, is typical (26–45). However, any claim that data recorded on magnetic strips or other computer storage media is semasiographic writing must be immediately complicated by the fact that the data is not typically visible to the human eye, and if it is visible then it is not typically meaningful. In short, whether graphic or phonetic in its basis, writing is always defined in terms of communication, and the inscriptions in computational storage media generally fail to communicate effectively to a human being absent the aid of a mechanical prosthesis. As Winfried Nöth notes in his *Handbook of Semiotics* (Bloomington: Indiana University Press, 1990), however, "In cybernetics and systems theory communication is often the interaction between any two entities. Thus, Klaus (1969) in his dictionary of cybernetics defines communication as 'the exchange of information between dynamic systems capable of receiving, storing, or transforming information.' This definition of communication also includes processes of interaction between machines" (170–171). Thus there is a clear tradition of including machine processes in the context of communications systems. Other discussions of semiotics come to rest on the distinction between communication and signification. Typically the former is associated with intentionality, whereas the latter category, signification, is broader and encompasses nonintentional signals (Nöth 172–173). Umberto Eco, however, reverses the two: "For Eco, any flow of information from a source to a destination is a process of communication, even the passage of a signal from machine to machine" (Nöth 172). And there is yet another possibility. Rather than writing in the orthodox sense of semasiography or glottography (also known as lexigraphy), computer data inscription is more akin to the forms of symbolic numeracy found in practices of record keeping. Yet Salomon argues persuasively not only that writing has its origins in record-keeping tasks (as is well known), but that "the development of a writing system is nothing other than the practical case-by-case solution of social tasks which produces an emergent new data registry system . . . the record-keeping art takes shape around the social problems it solves" (28). In short, he posits a continuum rather than a break between writing and recording.

In addition to whatever semiotic/cybernetic arguments one wants to entertain about the status of data inscriptions as writing, it is equally useful and important to juxtapose the prevalence of writing-related terminology in computer science practice. Even lay users routinely speak of reading and writing to and from a disk, and a disk drive includes an electromechanical instrument known as the read/write head. Are these metaphors or literal descriptors? Or else consider that "words" were a basic organizational unit in early data stores; a "word," according to Ceruzzi, was either an 11-digit number plus a sign, a string of 12 characters, or two 6-character instructions (23). If a computer scientist can speak quite literally about "writing" "words" without recourse to either glottographic or semasiographic writing systems, then quotation marks as I just used them are inappropriate and the definition of writing should be expanded to accommodate the phrase without them—just as the definition of "language" has expanded to include programming languages as well as natural languages without necessarily positing equivalences between them.

There is, of course, an obvious sense in which these marks and traces are meant to be machine readable and are here only incidentally revealed to the human visual field as the result of a rehearsed procedure. But it would be a mistake to think that the boundary between human and machine reading is always absolute or inflexible. The history of codes reveals a continuum rather than an absolute rupture between human and machine reading.[7] Early telegraph operators quickly learned to decode messages by listening to the sound of the receiver's mechanism rather than referring to the printed Morse it outputted. UPC symbols are legible to a trained observer. Punch cards can be manually deciphered and perhaps even more tellingly, their proper interpretation can be disputed, as vividly demonstrated during the 2000 election controversy in the United States. A computer forensics expert can visually inspect the patterns of magnetic tracks on a diskette treated in the same manner as the Metro card above and locate the starting points for the different data sectors. Still, all of these examples are admittedly specialized. Little wonder then that electronic writing's first generation of theorists turned their gaze toward the illuminated screen rather than the inscrutable disk. "[T]he simple, and possibly profound, truth," writes Xerox document scientist David Levy, "is that you can't see bits. You can't see them, you can't hear them, you can't touch or smell them. They are completely inaccessible to the human senses." Jay David Bolter puts it this way: "If you hold a magnetic or optical disk up to the light, you will not see text at all. At best you will see the circular tracks into which the data is organized, and these tracks mean nothing to the human eye."[8] The cathode ray tube was the implicit, and often explicit, starting point for most discussions of electronic textuality because it was only as bit-mapped fonts on the screen that electronic letterforms became recognizable as writing.[9] Critics such as Richard Lanham were quick to comment on the implications of desktop publishing and digital typography, noting that the creative control afforded by the font libraries and clip art galleries at every user's fingertips contributed to the breakdown of traditional distinctions between reader and writer while dramatizing the malleability of words and images in a digital setting.[10]

Nick Montfort has coined the term "screen essentialism" to refer to the prevailing bias in new media studies toward display technologies that would have been unknown to most computer users before the mid-1970s (the teletype being the then-dominant output device). One result, as Montfort discusses, is that an essential dimension of the materiality of early electronic literary productions like ELIZA and ADVENTURE is elided, since these works were historically experienced as printed texts on rolls of paper rather than as characters on video screens.[11] Thus one does not always need to look at screens to study new media, or to learn useful things about the textual practices that accumulate in and around computation. In their book *The Myth of the Paperless Office*, Abigail J. Sellen and Richard H. R. Harper employ J. J. Gibson's concept of affordances to evoke the raw, literal, physical materiality of different kinds of objects and media, especially paper: "The physical properties

[7] For an overview, see Charles Petzold, *Code: The Hidden Language of Computer Hardware and Software* (Redmond: Microsoft Press, 1999), especially chapters 1–10.

[8] David M. Levy, *Scrolling Forward: Making Sense of Documents in the Digital Age* (New York: Arcade Publishing, 2001), 138, and Jay David Bolter, *Writing Space: The Computer, Hypertext, and the History of Writing* (Hillsdale, NJ: Lawrence Erlbaum, 1991), 42.

[9] In fact, the early limitations of the Macintosh with its low-resolution VDT display were quickly enlisted by type designers such as *Émigré*'s Zuzana Licko—who began working seriously with the Mac within weeks of its debut—to provide the basic components of an electronic graphical identity. Licko says of this process: "I started my venture with bitmap type designs, created for the coarse resolutions of the computer screen and dot matrix printer. The challenge was that because the early computers were so limited in what they could do you really had to design something special. . . . it was physically impossible to adapt 8-point Goudy Old Style to 72 dots to the inch. In the end you couldn't tell Goudy Old Style from Times New Roman or any other serif text face. . . . It is impossible to transfer typefaces between technologies without alterations because each medium has its peculiar qualities and thus requires unique designs." See Rudy VanderLans and Zuzana Licko with Mary E. Gray, *Émigré (the Book): Graphic Design into the Digital Realm* (New York: Van Nostrand Reinhold, 1993), 18 and 23. What began as a material limitation in the medium's underlying hardware and display technologies was quickly accepted, adopted, and adapted as an integral aspect of the medium's aesthetic identity, an identity which has remained iconically intact and recognizable (think jaggies) even today, long after the technological base has shifted beyond the crude conditions Licko describes above.

[10] See Richard Lanham, *The Electronic Word: Democracy, Technology, and the Arts* (Chicago: University of Chicago Press, 1993), 3–28.

[11] Nick Montfort, "Continuous Paper: The Early Materiality and Workings of Electronic Literature": http://nickm.com/writing/essays/continuous_paper_mla.html.

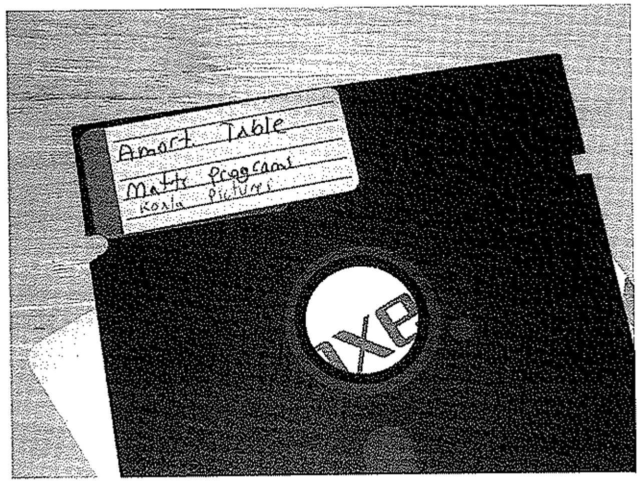


Figure 1.2 Affordances of a $5\frac{1}{4}$-inch floppy. Photograph by the author.

of paper (its being thin, light, porous, opaque, flexible, and so on) afford many different human actions, such as grasping, carrying, manipulating, folding, and in combination with a marking tool, writing on."[12] For Sellen and Harper (a cognitive psychologist and a technologist respectively), affordances are all about possibilities for action, which determine how human beings interact with the physical things in their environment. Computer storage media also have their affordances, but as storage in general has become more capacious and less immediately tangible it is easy to overlook them. (USB thumb drives are perhaps the best example of a recent storage innovation whose affordances have changed the way we interact with data: they are small and lightweight but also rugged, more like an accessory or gear than the flatter, flimsier profile of media like CD-ROMs or disks, whose vulnerable surfaces must be sheltered.) Attention to the affordances of various kinds of storage media can reveal much about computing in different contexts, allowing us to reconstruct salient aspects of now-obsolete systems and the human practices that attended them.

This is a $5\frac{1}{4}$-inch "floppy" disk, a personal relic from my teenage years (figure 1.2). The Apple II computer with which this disk is compatible has no hard drive. A program is loaded by inserting the disk in the external drive and booting the machine. In practical terms, this meant first retrieving the program by going to one's collection of disks and rummaging through them—perhaps they were kept in a shoebox, or stacked in a pile next to the computer, or in one of the many dedicated media containers marketed to the home computer enthusiast. Consider the contrast in affordances to a file system mounted on a hard drive: here you located the program you wanted by reading a printed or handwritten label, browsing like you would record albums or manila file folders, not by clicking on an icon. Written labels were therefore indispensable, their legible text standing in implicit counterpoint to the machine-readable markings on the magnetic surface sheathed within the plastic envelope to which the label was affixed. The label on this particular disk—handwritten, and placed over the manufacturer's label—indicates three different items: "Amort. Table" (amortization table, a tool used by accountants—why this would have been of interest to my fourteen-year-old self is beyond me now), "Matts [sic] Programs" (some programs I had written in BASIC), and (penned with an obviously different ink) "Koala Pictures" (Koala was a drawing utility—these were its data files). So on the same diskette we have, commingled, a freeware business application, programs I had written by myself and for myself, and the data files created by a commercial software package. The latter appears to have been added at a later date. This alone—the heterogeneous nature of the content, its incremental consignment to the disk—tells us something about the culture of personal computing at the time (which was clearly different from the affordances even of CD-Rs today.) In addition, we can see that a portion of the disk envelope alongside the label has been crudely cut away; in fact, it was attacked with a hole puncher. By mimicking the professionally cut write tab in a symmetrical location I double-sided the disk, signaling to the drive mechanism that the reverse side was also available for data inscription (to access the reverse side you would place the disk in the drive upside down). This was a common trick, and one I was quick to appreciate once I learned that disks had a finite capacity, that more disks were what enabled me to expand my software collection (primarily games), and that the money for buying new blank disks was to come out of my allowance. In this instance I can surmise that I double-sided this disk at some point well after its initial purchase in order to store the Koala picture files I was then generating with that piece of software (itself stored on a

[12] Abigail J. Sellen and Richard H. R. Harper, *The Myth of the Paperless Office* (Cambridge: MIT Press, 2001), 12.

separate disk), hastily inking in an addition to the label which would allow me to locate them.

I am belaboring these details to make the point that as a teenage computer user I had unself-consciously worked with storage media whose material qualities were very particular but which differ markedly from what would be the norm today. Since even routine chores like disk defragmentation are performed far less frequently on the current generation of hard drives, storage has become ever more of an abstraction, defined only by a volume letter (on most Windows machines, "C"), a graphic hard drive icon, or a pie chart visualization of space remaining. Greater and greater storage capacity will only serve to further dematerialize the media as their finite physical boundaries slip past the point of any practical concern. Compare this to the kind of information preserved on the manufacturer's labels of floppy disks from the 1980s, emblems of a bygone inscriptive regimen: "Reliable and Durable" promises the label on one brand, "48 TPI" (tracks per inch) specifies another, and "Double-Sided/Double-Density Soft-Sectored With Hub Ring" declares a third. This strange and alien cant was perfectly intelligible to me and millions of other home computer users, not because we were hackers or übergeeks, but because these specs defined the functional limits of what we could and could not do with a given piece of media in practical and palpable ways—in other words, its affordances.

A further contrast between screen essentialism and inscription or storage media is warranted, I believe, by the current state of new media studies in which the graphical user interface is often uncritically accepted as the ground zero of the user's experience. "We look through the interface unaware," writes Michael Heim in his *Metaphysics of Virtual Reality*, "as we peer through an electronic network where our symbols—words, data, simulations—come under precise control, where things appear with startling clarity. So entrancing are these symbols that we forget ourselves, forget where we are. We forget ourselves as we evolve into our fabricated worlds. With our faces up against it, the interface is hard to see."[13] Heim's experience here speaks powerfully to a technological sublime, a simultaneous ecstasy and oblivion immanent in our encounters with the virtual. But this "metaphysics," to use Heim's word (a metaphysics conceived, one suspects, amid the vertigo of Gibson's city lights

[13] Michael Heim, *The Metaphysics of Virtual Reality* (New York: Oxford University Press, 1993), 79–80.

receding), is not finally symbolic (note that word's repetition in his text) but instead embedded within real-world technologies of production and design. Robert Markley, writing partly in direct response to Heim (in an essay that should be better known than it is), offers a prescient brute force disassembling of screen essentialism:

> To ask what is on the other side of the computer screen is, in my mind, a crucial step in dissenting from this consensual hallucination. Behind the screen of my laptop lie silicon chips, a battery, microprocessors, and even what seem to be a few old-fashioned screws. It runs (now rather dated) software programs engineered originally in California and Utah. My access to the presumptive world behind the screen carries with it an effaced history of labor, of people building machines to design and to build even more sophisticated hardware and software. (77)[14]

"The imaginary realm of cyberspace," Markley concludes, ". . . is a fantasy based on the denial of ecology and labor, a dream that is also an apology for the socioeconomic power to bring together sophisticated technologies" (77). Markley's account is one of the few from this era to explicitly juxtapose the gaze of the end user with the unseen workers' hands—here literally screened from view—which are busy turning old-fashioned screws.[15]

Yet even in Markley's resolutely anti-essentialist hands, the screen still seems to slip into a synecdoche for "the computer" as a whole. What I have been attempting to accumulate here are thus a set of alternative access points for the study of computing, access points that bring storage, inscription, and engineering into the visible purview of what we think of as new media. But how did screens come to so obscure our view in the first place?

[14] Robert Markley, "Boundaries: Mathematics, Alienation, and the Metaphysics of Cyberspace," in *Virtual Reality and its Discontents*, ed. Robert Markley (Baltimore: The Johns Hopkins University Press, 1996): 55–77.

[15] Precisely this dynamic is explored in the 1994 corporate thriller *Disclosure*. The high-tech company that is the setting for the sexual harassment charges driving the film is showcasing a fanciful, fully immersive virtual reality environment with stunning visuals (or what passed for them at the time) as its next generation technology. Its current product, however, and the focal point for the plot, is a line of high-speed compact disk drives. As my colleague Katie King points out, here storage is made visible through the plot's attention to the manufacturing process and the associated industrial espionage. The climax of the film occurs when evidence of tampering with the drives is "disclosed"—on a big-screen TV—at a shareholders meeting.

A Medial Ideology

Jerome McGann has used the phrase "Romantic ideology" to describe the manner in which modern literary criticism of the Romantic poets has been characterized by "an uncritical absorption in Romanticism's own self-representations."[16] I believe electronic textual theory has labored under similar uncritical absorptions of the medium's self- or seemingly self-evident representations. While often precisely Romantic in their celebration of the fragile half-life of the digital, the "ideology" I want to delineate below is perhaps better thought of as *medial*—that is, one that substitutes popular representations of a medium, socially constructed and culturally activated to perform specific kinds of work, for a more comprehensive treatment of the material particulars of a given technology.

This tendency is already full-blown in Arthur C. Clarke's 1986 short story "The Steam-Powered Word Processor," which narrates the fictitious history of the Reverend Charles Cabbage (obviously a stand-in for the historical Babbage), vicar of the tiny church in Far Tottering, Sussex.[17] As Clarke tells it, the Reverend, weary of his obligation to produce varying sermons on the same theme twice a week, 104 times a year, contrives to build a device for automating the composition process. The Word Loom is envisioned as a combinatory machine for the manipulation of sentences (which it takes as the basic combinatory unit), and it is to be capable of outputting hard copy for the Reverend's use by way of something like a Linotype process.

The machine's database is the Bible and Cruden's Concordance, punched onto cards "at negligible expense, by the aged ladies of the Far Tottering Home for Relics of Decayed Gentlefolk" (932). Having solved the problem of data entry and beat Herman Hollerith to the punch (as it were), Cabbage (who we also learn enjoys a correspondence with the aging Michael Faraday) proceeds to other aspects of his design. The church's pipe organ becomes his chief inspiration: "He was convinced that an assembly of pneumatic tubes, valves, and pumps could control all the operations of his projected Word Loom" (932). But the reader quickly intimates that this ambitious enterprise is doomed from the start, Cabbage's novel solutions to problems of input, output, processing, and storage notwithstanding. On the day of its first and only public trial something goes awry—"Somewhere, in the depths of the immense apparatus, something broke" (933)—and the Word Loom is rent to pieces in a maelstrom of imploding machinery. All that survives today are "two or three gearwheels" and "what appears to be a pneumatic valve" in the possession of the Far Tottering Historical Society (932); that and, deep inside the British Museum, bound in a volume entitled *Sermons in Steam*, a single machine-generated page, badly printed and riddled with typographical errors. It is either a clever fake (we are told) or else it is "the only surviving production of perhaps the most remarkable—and misguided—technological effort of the Victorian Age" (934).

[16] Jerome McGann, *The Romantic Ideology* (Chicago: University of Chicago Press, 1983), 1.

[17] In *The Collected Stories of Arthur C. Clarke* (New York: Tor, 2000), 930–934. All pages references are to this edition. Originally published in *Analog*, January 1986.

Of course no reader accepts this conceit, and it doesn't matter; the lesson is all about the folly of seeking to embed digital behaviors in an industrial engine. The Word Loom, which was to "weave thoughts the way Jacquard wove tapestries" (932; note the paraphrase of Ada, Countess Lovelace), could only succeed with the aid of new forces harnessed by Faraday in his work on electromagnetic energy, not the brass gearwheels of a Victorian mechanism. It is Cabbage's hapless lot to seek to pour new wine into a very old bottle. We, savvy readers and beneficiaries of the technology of word processing ourselves, are in on the joke and therefore understand that the story serves chiefly to underscore the radical break between electronic writing and earlier forms of textuality—a familiar and comfortable enough homily for old Cabbage to deliver in the end.

We can pick up the thread of the Word Loom just two years later when Umberto Eco, in his novel *Foucault's Pendulum*, contrives the kabbalistically named word processor Abulafia as the embodiment of a "totally spiritual machine" (24):

> If you write with a goose quill you scratch the sweaty pages and keep stopping to dip for ink. Your thoughts go too fast for your aching wrist. If you type, the letters cluster together, and again you must go at the pokey pace of the mechanism, not the speed of your synapses. But with [Abulafia] your fingers dream, your mind brushes the keyboard, you are borne on golden pinions, at last you confront the light of critical reason with the happiness of a first encounter. (24–25)[18]

[18] Umberto Eco, *Foucault's Pendulum*, Trans. William Weaver (London: Picador, 1989).

Already we can glimpse the particulars of our medial ideology: "Our best machines are made of sunshine; they are all light and clean because they are nothing but signals, electromagnetic waves, a section of a spectrum . . ." (153) wrote Donna Haraway in her famous "Cyborg Manifesto" (1985; that her words were also ironic strengthens rather than diminishes their medial impact).[19] Industry leaders may have grasped the appeal of this ideology even earlier than fiction writers or academicians. In 1982, four Bay-area entrepreneurs cofounded a new company devoted to network enterprise computing. They called it Sun.

By the mid-1980s, the digital sphere had assumed visual and material form as a definable and datable set of aesthetic practices; a recognizable spectrum of tropes, icons, and graphic conventions. This is the backdrop for the medial ideology I am describing. At stake is not whether such conventions for representing digital phenomena are accurate or correct to the formal ontology of information in an absolute sense, but rather the important fact that Western consumer culture had succeeded in evolving sophisticated and compelling conceits for depicting information as an essence unto itself, or more properly, information as a synthetic (at times even haptic) commodity. That the cyberspaces of both *Neuromancer* and *Tron* (as well as other cyberpunk productions such as the short stories in Gibson's "Burning Chrome" anthology or even the 1981 animated feature *Heavy Metal*) are artificial alloys derived of complex cultural skeins may seem an elementary point, but it is one that was often lost in the face of popular enthusiasms for virtual phenomena. A touchstone would be Michael Benedikt's anthology *Cyberspace: First Steps*, published in 1991, which collected the writings of many of the so-called digerati, the loose clique of artists and technologists who had emerged over the course of the previous decade.[20] Prefaced by Gibson, the volume contains fifteen essays, notable today for how literally some of them read Gibson's novels as starting points for actual research agendas in interface design and related fields. All of the essays oscillate between tacit recognition of the preliminary and tentative status of the actual technologies on the one hand, and a willingness to talk about cyberspace as though it were already an observable phenomenon on the other. Some contributors simply choose not to acknowledge this as an issue; Marcos Novak, for example, does not hesitate to inform us that "The function of [cyberspace synthesizers] is to receive a minimal description of the cyberspace, coded and compressed, and from it to render a visualization of that space for the user to navigate with" (233). This is a reasonable enough description of the graphical Web browsers that would soon emerge, but one wonders if Novak didn't have something more fanciful in mind. In either case, he is assuming cyberspace is subject matter evocative enough for the reader to suspend disbelief and to benefit from a putatively sober description of a technology that does not yet exist. Other contributors are more circumspect, such as David Tomas, who asserts the following: "Although cyberspace has been popularized by Gibson's books, it is neither a pure 'pop' phenomenon nor a simple technological artifact, but rather a powerful, collective, mnemonic technology that promises to have an important, if not revolutionary, impact on the future compositions of human identities and cultures" (31–32). This appears a balanced assessment, yet it is clear that when Tomas talks about cyberspace as a "technology" he cannot mean technology in the sense of any specific hardware or software implementation—a meaning he hastens to jettison by preceding his reference to a "technological artifact" with the qualifier "simple" and by placing the whole of the phrase in parallel with the equally ineffectual notion of cyberspace as a "pure 'pop' phenomenon." Cyberspace, as it is invoked here, can only be a technology in the sense that the word itself—or more precisely, the idea of cyberspace—mimics the behavior of certain material technologies, functioning as a "powerful, collective, mnemonic"—or in other words, as a shorthand for a whole range of communicative agendas given depth and form by a shared aesthetic. This is what was reflected in *Wired*'s Teflon sheen when the magazine, self-consciously designed to *look* like information, began publishing in 1993.[21]

At the core of a medial ideology of electronic text is the notion that in place of inscription, mechanism, sweat of the brow (or its mechanical equivalent steam), and cramp of the hand, there is light, reason, and energy unleashed in the electric empyrean. Yet Clarke, Eco, and others among the first to write about word processors, both on and with their own home computers (such as Michael Joyce, whose experience we will examine in chapter 4), were not

[19] In Donna J. Haraway, *Simians, Cyborgs, and Women: The Reinvention of Nature* (New York: Routledge, 1991), 149–181.

[20] Michael Benedikt, ed. *Cyberspace: First Steps* (Cambridge: MIT Press, 1991).

[21] Note also the remarkable visual consistency to the maps and renditions of cyberspace, both scientific and imaginative, on display at Martin Dodge's *Atlas of Cyberspaces*: http://www.cybergeography.org/atlas/atlas.html.

simply deluded or wrong. Indeed, what was *new* about the technology was precisely that it succeeded so completely in rendering the workaday labor of textual production *functionally* immaterial. Michael Heim successfully articulated an early intuition as to why. In his *Electric Language: A Philosophical Discussion of Word Processing* (1987), which precedes now better-known books by Richard Lanham, George Landow, and Jay David Bolter, he borrows the term "system opacity" from John Seely-Brown:

> The types of physical cues that naturally help a user make sense out of mechanical movements and mechanical connections are simply not available in the electronic element. There are far more clues to the underlying structural processes for the person riding a bicycle than there are for the person writing on a computer screen. Physical signs of the ongoing process, the way that responses of the person are integrated into the operation of the system, the source of occasional blunders and delays, all these are hidden beneath the surface of the activity of digital writing. No pulleys, springs, wheels, or levers are visible; no moving carriage returns indicate what the user's action is accomplishing and how that action is related to the end product; and there is no bottle of white-out paint complete with miniature touch-up brush to betoken the industrial chore of correcting errors by imposing one material substance over another. The writer has no choice but to remain on the surface of the system underpinning the symbols.[22]

System opacity or black box: what was implicit in Eco's paen to dreaming fingers borne aloft becomes explicit in Heim: "Yet, in order to achieve such automation, *writing has to be removed from the element of inscription* and placed in an electronic element" (136; emphasis added). It is not that Heim is oblivious to the operations of his disks and storage peripherals; on the contrary, he insists that some basic understanding of how the computer stores and retrieves information is essential for even a novice user, else they will be duped by watching text scroll off the edge of the screen. But Heim conceives of this understanding pragmatically, as "a set of metaphors for making operational guesses at the underlying structure" (133)—not in terms of specific technologies. Thus in Heim's example, a user might imagine that two different versions of a document are saved in two different "places" on the hard drive (133–134). As

[22] Michael Heim, *Electric Language: A Philosophical Study of Word Processing*, 2nd edition (New Haven: Yale University Press, 1999), 131–132.

Heim himself is careful to point out, "[t]his insight requires no awareness of the File Allocation Table (FAT) or the bits set for file identifiers on the level of machine-language bytes, nor does it require awareness of the tracking system on the disk drive" (134). Electronic writing thus becomes a friction-free "information flow" (133), an essentially symbolic rather than inscriptive exchange among a set of operational metaphors and the "electronic elements" on the screen. Later in this book, I will term this symbolic exchange formal materiality.

Meanwhile the academy had begun a conversation in earnest. In a chapter titled "Derrida and Electronic Writing" from the 1990 book *The Mode of Information*, Mark Poster described electronic textuality this way:

> Compared to the pen, the typewriter or the printing press, the computer dematerializes the written trace. As inputs are made to the computer through the keyboard, pixels of phosphor are illuminated on the screen, pixels that are formed into letters. Since these letters are no more than representations of ASCII codes contained in Random Access Memory, they are alterable practically at the speed of light. The writer encounters his or her words in a form that is evanescent, instantly transformable, in short, immaterial.[23]

And one year later, in what may be the most influential critical study to emerge from this era, Jay David Bolter in *Writing Space* (1991):

> Electronic text is the first text in which the elements of meaning, of structure, and of visual display are fundamentally unstable. . . . This restlessness is inherent in a technology that records information by collecting for fractions of a second evanescent electrons at tiny junctures of silicon and metal. All information, all data in the computer world is a kind of controlled movement, and so the natural inclination of computer writing is to change, to grow, and finally to disappear.[24]

Or George P. Landow and Paul Delany, also in 1991:

[23] Mark Poster, *The Mode of Information: Poststructuralism and Social Contexts* (Chicago: University of Chicago Press, 1990), 111.

[24] Jay David Bolter, *Writing Space: The Computer, Hypertext, and the History of Writing* (Hillsdale, NJ: Lawrence Erlbaum, 1991), 31.

> So long as the text was married to a physical media, readers and writers took for granted three crucial attributes: that the text was *linear*, *bounded*, and *fixed*. Generations of scholars and authors internalized these qualities as the rules of thought, and they have pervasive social consequences. We can define *Hypertext* as the use of the computer to transcend the linear, bounded and fixed qualities of the traditional written text.[25]

The preceding accounts are not without nuance. In Poster, for example, who is in turn reminiscent of Eco, electronic writing is presented in relative and not absolute terms as compared to the pen or typewriter or printing press, and he specifies the computer's RAM as the site of the written act. Yet this very specificity makes the absence of storage all the more telling. Landow and Delany, interestingly, define hypertext first and foremost as the *use* of a technology, rather than a technology itself. Yet we still find a clear opposition between the expressive largesse electronic environments promise and textual stability which is relegated to "physical media," the presumed baggage of the Gutenberg galaxy.

We can continue to accumulate examples of a medial ideology, year to year. "[W]ith electronic text we are always painting, each screen unreasonably washing away what was and replacing it with itself," suggested Michael Joyce in 1992.[26] N. Katherine Hayles, in a widely read essay titled "Virtual Bodies and Flickering Signifiers" published a year later, puts it this way: "Working at the computer screen, I cannot read unaided the magnetic markers that physically embody the information within the computer, but I am acutely aware of the patterns of blinking lights that comprise the text in its screen format" (260).[27] Hayles, of course, fully understands the internal complexity of the symbolic transaction she is alluding to, noting elsewhere that the screen's "flickering signifiers" (as she calls them, after Lacan) originate as magnetic traces on a disk, which are then interpolated through binary machine language, compiler, application software, and so forth. The basic thesis Hayles goes on to develop, that signification in electronic environments involves "a flexible chain of markers bound together by the arbitrary relations specified by the relevant codes" (264), effectively captures what most users experience as the basic phenomenological difference between analog and digital media (whether backspacing to correct a typing error or brightening an image in Photoshop). Digital signification, in this model, consists in an open-ended symbiotic exchange (or feedback loop) between computation and representation.

I believe that the close temporal proximity and occasional outright overlap in the passages I have been collecting point not to a transparent and self-sufficient account of the ontology of the medium itself, but rather to the emerging contours of a medial ideology. We see this best if we examine the dominant tropes and rhetorical markers. Speed and light (or lightning) are paramount: Heim's "electric language," Eco's "golden pinions" and "light of critical reason," Poster's "pixels of phosphor" and "speed of light," Bolter's "fractions of a second" and "evanescent electrons," Hayles's "patterns of blinking lights" and "flickering" signifiers, even Joyce's spontaneous "painting." This is consistent with the tropes that had already emerged in narrative and cinematic science fiction: the luminous, ray-traced aesthetic of *Tron*, or William Gibson's prose passages describing cyberspace in terms of "lines of light" and "city lights receding."

Recall Heim, for whom symbolic automation demands medial liberation: writing had to be "removed from the element of inscription" and placed in electronic form. With letterforms now "instantly transformable," we reach for the chiasmus of "fingers dreaming" and a "mind brushing the keyboard" (shades of Gibson's anti-hero Case and his Ono-Sendai cyberspace deck). We invoke adjectives like "flickering," "restless," "flexible," and their ultimate apotheosis, "immaterial." This medial ideology is precisely the same aesthetic to which *Agrippa* spoke so powerfully in 1992, with its then-contemporary meme of the viral, self-consuming text and the disappearing book. Screen essentialism becomes a logical consequence of a medial ideology that shuns the inscriptive act.

There are a number of important respects in which the theoretical debate has advanced considerably since the first half of the 1990s. The appeal to high poststructuralism forming the backdrop of many of the early accounts I have referenced has been abandoned, or at least its influence diluted. Espen Aarseth's *Cybertext: Perspectives on Ergodic Literature* (1997) emerged as the first major attempt to examine screen-level effects from the vantage point of their

[25] Paul Delany and George P. Landow, eds. *Hypermedia and Literary Studies* (Cambridge: MIT Press, 1991), 3. Emphases in original.

[26] Michael Joyce, *Of Two Minds: Hypertext Pedagogy and Poetics* (Ann Arbor: University of Michigan Press, 1995), 232.

[27] N. Katherine Hayles, "Virtual Bodies and Flickering Signifiers," in *Electronic Culture: Technology and Visual Representation*, ed. Timothy Druckrey (New York: Aperture, 1996), 259–277.

interaction with a text's underlying formal processes, leading to, among much else, a widening of the general purview of the field of electronic textual studies (video games, old school interactive fiction, and printed text machines such as *Choose Your Own Adventure* novels, the ancient *I Ching*, and Oulipean productions like Raymond Queneau's *Cent Mille Milliards de poèmes*).[28] N. Katherine Hayles, meanwhile, continued to refine her critical positions and has advanced what is probably the most extensive argument for materiality and embodiment against the backdrop of three recent books and a series of accomplished, media-specific close readings of printed and electronic texts alike.[29] In his widely read *The Language of New Media* (2001), Lev Manovich developed an extensive formal account of new media influenced by his background in film, but the book's most important contribution may yet be its advocacy of software studies, the serious study of specific software technologies as both historical and computational artifacts, a call that has also been taken up by Matthew Fuller.[30] Alan Liu's *The Laws of Cool* (2004), a majestic book whose fundamental frame of reference is as much the cubicle as the screen, locates its potential for a "future literary" within the cool enclaves of the Web nurtured by digital artists and corporate knowledge workers alike.[31] At an even more general level, new media studies has seen essential critical work on the politics of race, class, and gender;[32] the expansion into brand new areas, notably the

[28] Espen Aarseth, *Cybertext: Perspectives on Ergodic Literature* (Baltimore: Johns Hopkins University Press, 1997).

[29] The three books are *How We Became Post-Human: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago: University of Chicago Press, 1999), *Writing Machines* (Cambridge: MIT Press, 2002), and *My Mother Was a Computer: Digital Subjects and Literary Texts* (Chicago: University of Chicago Press, 2005).

[30] Lev Manovich, *The Language of New Media* (Cambridge: MIT Press, 2001); see also Fuller's *Behind the Blip: Essays on the Culture of Software* (Brooklyn, NY: Autonomedia, 2003).

[31] Alan Liu, *The Laws of Cool: Knowledge Work and the Culture of Information* (Chicago: University of Chicago Press, 2004).

[32] For example, Anne Balsamo, *Technologies of the Gendered Body: Reading Cyborg Women* (Durham: Duke University Press, 1996); Lisa Nakamura, *Cybertypes: Race, Ethnicity, and Identity on the Internet* (London: Routledge, 2002); *Reload: Rethinking Women + Cyberculture*, eds. Mary Flanagan and Austin Booth (Cambridge: MIT Press, 2002); and Geert Lovink's *Dark Fiber: Tracking Critical Internet Culture* (Cambridge: MIT Press, 2003), and *Uncanny Networks: Dialogues with the Virtual Intelligentsia* (Cambridge: MIT Press, 2004).

white-hot field of ludology;[33] and the recent availability of essential basic reference tools such as *The New Media Reader* and *Information Arts*.[34] Yet for all of this activity, my argument is ultimately that we remain very much in the grip of a medial ideology, with many of the plain truths about the fundamental nature of electronic writing apparently unknown at a simple factual level, or else overlooked or their significance obscured.

In the next section, I introduce the field of computer forensics as a counterpoint. At the applied level, computer forensics depends upon the behaviors and physical properties of various computational storage media. For the theoretical observer, it is bracing to watch the forensic investigator run up against many of the same issues that have driven commentators in the realm of cultural and critical theory. While digital evidence can be instantly deleted it can often be just as easily recovered; while digital evidence can be copied perfectly (what we like to call a simulacrum), it can also be copied imperfectly, and in fact care must be taken lest it be copied incompletely; while digital evidence can be tampered with, it can also be stabilized and encrypted; while digital evidence can be faked, it can also be signed and algorithmically authenticated. In a very different climate from the anxieties of the academy discussed above, the forensic investigator employs a set of field procedures designed to establish an order of volatility for all of the evidence within his or her purview, clinically delineating the relative vulnerability and stability of data at many different points in a system's internal architecture. The irony is that while the protected internal environment of the hard drive is built to exclude the hairs, fibers, and other minute particulars of traditional forensic science, the platter inexorably yields up its own unique kind of physical evidence.

Computer Forensics

According to one definition, computer forensics consists in "the preservation, identification, extraction, documentation, and interpretation of computer

[33] See Nick Montfort, *Twisty Little Passages: An Approach to Interactive Fiction* (Cambridge: MIT Press, 2003); Noah Wardrip-Fruin and Pat Harrigan, eds. *First Person: New Media as Story, Performance, and Game* (Cambridge: MIT Press, 2004) and Jasper Juul, *Half-Real: Video Games Between Real Rules and Fictional Worlds* (Cambridge: MIT Press, 2005).

[34] Noah Wardrip-Fruin and Nick Montfort, eds., *The New Media Reader*, (Cambridge: MIT Press, 2003) and Stephen Wilson, *Information Arts: Intersections of Art, Science, and Technology* (Cambridge: MIT Press, 2003).

data."[35] Other definitions also emphasize the data's status as physical evidence.[36] Computer forensics is the activity of recovering or retrieving electronic data, analyzing and interpreting it for its evidentiary value, and preserving the integrity of the data such that it is (potentially) admissible in a legal setting. At a practical level this means working with hard drives and other storage media in the field and in controlled laboratory settings to locate files, metadata, or fragments of files that someone may or may not have taken active steps to expunge, and creating the conditions necessary to ensure that the data has not been tampered with in the process of its recovery or analysis. Precedents, case law, and statutes date back to the late 1970s, but computer forensics has really only emerged as a professional specialization in the last five to ten years. There are now a number of textbooks on the shelves, growing numbers of specialized software tools (some retailing for many thousands of dollars), specialized hardware like self-contained drive imaging units, and elite corporate training programs. The most advanced computer forensics, however, undoubtedly happens not in commercial settings but through government agencies like the FBI, the NSA, the National Center for Computer Security, and the U.S. Department of Defense's Cyber Crime Center. Computer forensics in fact transcends the investigation of so-called cyber crime (such as identity theft) to claim a much broader purview. As the textbooks unfailingly point out, the search and seizure of digital evidence has become a routine part of many criminal investigations. The BTK Killer is an example of one recent high-profile case where computer forensics furnished the major break: Dennis Rader was identified and apprehended after residual data on a floppy disk he sent to a local TV station allowed authorities to pinpoint the computer where the disk's files had originally been created.[37] Likewise, popular awareness of computer forensics has grown through the popularity of television drama and genre fiction. Nor is its purview limited to desktop computers and laptops. One government lab I visited prides itself on its ability to retrieve and analyze data from the full spectrum of electronic devices: pagers, cell phones, personal digital assistants, GPS units, digital watches, game consoles, digital cameras, magnetic access cards, programmable appliances, automotive chips, and more.[38]

The most relevant forensic science precedent for computer forensics is the field of questioned document examination, which dates back to the end of the nineteenth century. It concerns itself with the physical evidence related to written and printed documents, especially handwriting attribution and the identification of forgeries. Questioned document examination also bears some resemblance to the academic pursuits of analytical and descriptive bibliography (which emerged in an organized fashion during roughly the same years), but the forensic enterprise subjects the myriad physical implements of writing and inscription to a degree of scrutiny that might give even a hardened bibliographer pause. (Handwriting identification as it is practiced by a forensic expert is not to be confused with graphology, the more dubious practice of deriving personality traits from the appearance of an individual's handwriting; this approach was explicitly rejected by the earliest texts on document examination, including Persifor Frazer's *Bibliotics, or the Study of Documents* [1894, itself now largely discredited] and William Hagan's *Disputed Handwriting* [also 1894].[39]) The questioned document examiner is almost

[35] Warren G. Kruse II and Jay G. Heiser, *Computer Forensics: Incident Response Essentials* (Boston: Addison-Wesley, 2002), 2.

[36] For example, Eoghan Casey, *Digital Evidence and Computer Crime* (Amsterdam: Academic Press, 2000), 4.

[37] As was widely reported in the media, for example, here: http://abcnews.go.com/WNT/story?id=539702&page=1.

[38] On April 21, 2005 I visited the Department of Defense's Defense Cyber Crime Center (DC3 for short), located in an anonymous office park near Baltimore/Washington International airport. I was received by Special Agent Jim Christy, the director of the lab, who began his career as a computer crime investigator with the Air Force in the 1970s. In conversation with Christy, a genial and engaging host, it became clear that the lab's two biggest challenges were the sheer volume of incoming data, and maintaining its accreditation in the face of constantly changing software. In terms of volume, a typical case might involve analyzing a hard drive with 80 GB of data. This is the equivalent of 1,360 file cabinets or 34 million pages of written material. Much of the investigation is automated, with technologies such as data mining and visualization playing a growing role. But the core of the analysis must be performed manually. In a typical year, DC3 might handle a volume of data that would fill 15 miles of physical file cabinets. This number will only increase. As for accreditation, each new generation of forensic software has to be vetted in order for any evidence it yields to be legally admissible in court. Keeping abreast of the accreditation process is a significant drain on staff and resources given how quickly software in the lab evolves through new versions and releases.

[39] Even Sherlock Holmes, author of a "little monograph" on the dating of documents, was not above the temptations of graphology: "'Look at his long letters,' he said. 'They hardly rise above the common herd. That *d* might be an *a*, and that *l* an *e*. Men of character always differentiate

always concerned with a document in its physical entirety. To quote Ordway Hilton, a contemporary authority:

> Not only must these examiners be able to identify handwriting, typewriting, and printed matter, but they must be able to distinguish forgery from genuineness, to analyze inks, papers, and other substances that are combined into documents, to reveal additions and substations in a document, and to restore or decipher erased or obliterated writing. When records produced by complex modern business machines are suspected of having been manipulated, document examiners may be among the first to be consulted.[40]

Many of these activities have explicit parallels in computer forensics as it is practiced today. Recovering erased data, authenticating digital documents, and identifying the source of an electronic object are all routine activities for the specialist. But while computer forensics may seem like a natural extension of the questioned document examiner's purview—Hilton's reference to "modern business machines" seems to point the way forward—in practice the two have remained separate domains. Questioned document examination's reference works, even very recent ones, tend to treat "computer documents" exclusively in terms of hard copy.

Both questioned document examination and computer forensics belong to a branch of forensic science known as "trace evidence," which owes its existence to the work of the French investigator Edmond Locard. Locard's famous Exchange Principle may be glossed as follows: "a cross-transfer of evidence takes place whenever a criminal comes into contact with a victim, an object, or a crime scene."[41] Locard, a professed admirer of Arthur Conan Doyle who worked out of a police laboratory in Lyons until his death in 1966, pioneered the study of hair, fibers, soil, glass, paint, and other small things forgotten,

their long letters, however illegibly they may write. There is vacillation in his *k*'s and self-esteem in his capitals' " (Sir Arthur Conan Doyle, *The Sign of Four*).

[40] Hilton, *Scientific Examination of Questioned Documents*, Revised Edition (New York: Elsevier, 1982), 4.

[41] Joe Nickell and John F. Fischer, *Crime Science: Methods of Forensic Detection* (Lexington: University Press of Kentucky, 1996), 10.

primarily through microscopic means. His life's work is the cornerstone of the stark dictum underlying contemporary forensic science: "Every contact leaves a trace." This is more, not less, true in the delicate reaches of computer systems. Much hacker and cracker lore is given over to the problem of covering one's "footsteps" when operating on a system uninvited; conversely, computer security often involves uncovering traces of suspicious activity inadvertently left behind in logs and system records. The 75-cent accounting error that kicks off Clifford Stoll's *The Cuckoo's Egg*, a hair-raising account of computer detective work that culminated in the seizure of several Eastern Bloc agents, is a classic example of Locard's Exchange Principle in a digital setting.[42]

Insights from computer forensics have the potential to overturn many of the chestnuts governing the critical conversation on new media and electronic textuality. Marcos Novak asserts the following, for example: "Everything that is written and transmitted via electronic media is erasable and ephemeral *unless* stored or reinscribed (emphasis added)."[43] My contention would be that the subordinating conjunction "unless" is called upon to do a great deal of unrealistic work. Practically speaking, most things that are written and transmitted via electronic media *are* stored and reinscribed. A simple e-mail message may leave a copy of itself on a half a dozen different servers and routers on the way to its destination, with the potential for further proliferation via mirrors and automated backup systems at each site. As storage costs continue to

[42] See Clifford Stoll, *The Cuckoo's Egg* (New York: Pocket Books, 1989). For a more literary expression of the exchange principle, there is this passage from Hari Kunzru's novel *Transmission* (New York: Dutton, 2004): "Whenever he entered and left the secure area, his bag was checked for storage media. As numerous laminated signs in the corridor pointed out, if a disk went into the [anti-virus] lab it did not come out again" (51).

[43] See Novak's passionate online essay "TransTerraForm" at http://www.krcf.org/krcfhome/PRINT/nonlocated/nlonline/nonMarcos.html. The sentence I quote is in the context of a broader argument that sets up a contrast between the supposed ephemerality of digital inscription and the literal inscription of microchips, which Novak describes as "immensely compactified books, active yet permanent, carved enduringly in silicon." Novak's essay is a meditation on the coming liquidity of inscription in chip design (of the sort now being realized by the MRAM technology I briefly discuss in the introduction). It is therefore unfortunate that the silicon chip is put forward as the primary site of inscription in the computer's architecture; the notion of "erasable, liquid" hardware configurations is a tantalizing one, but Novak misses the fundamental sense in which inscription is the essence of computer storage media, which is presented in the essay only as the site of the absent trace.

plummet, the trend will no doubt be to save more and more data so that the variety of ephemera routinely written to disk becomes ever more granular. Likewise, even the popular myth that RAM is always absolutely volatile, gone forever at the flip of a switch, proves false; there are at least experimental techniques for recovering data from RAM semiconductor memory.[44] While it may be technically possible to create the conditions in which electronic writing can subsist without inscription and therefore vanish without a trace, those conditions are not the medium's norm but the special case, artificially induced by an expert with the resources, skill, and motive to defeat an expert investigator.[45] For the remainder of this section I want to focus on three specific sets of observations from computer forensics in order to challenge some of the common assumptions about electronic textuality that characterize what I have been calling the medial ideology. They are as follows: that electronic text is hopelessly ephemeral, that it is infinitely fungible or self-identical, and that it is fluid or infinitely malleable.

Ephemerality Lay users often know that when they delete a file from their trash or recycle bin it is not immediately expunged from their hard drive. What happens instead is that the file's entry in the disk's master index (on Windows machines called the File Allocation Table or FAT, which we will discuss in more detail in the next chapter) is flagged as space now available for reuse. The original information may yet persist for some time before the

[44] See, for example, Peter Gutmann's discussion in "Data Remanence in Semiconductor Devices," http://www.cypherpunks.to/~peter/usenix01.pdf.

[45] Some readers may immediately think of dev/null, the file on the UNIX operating system used as a "bit bucket" because it is not attached to a disk or any other physical hardware device. Writing to dev/null consigns data to virtual oblivion, but while the existence of dev/null is common knowledge piping data to it is not the same as performing a secure deletion of a previously saved file. A further case in point, demonstrating the extremes one must go to in order to avoid leaving trace evidence, is a net art project by Greg Sidal. His "Illicit Images" consists of a set of color ink jet renderings of commercial image files originally harvested from the Web using automated collection routines routed through anonymous redirection services (to avoid leaving traces in access logs). These images are then encrypted so as to destroy all semblance of their previous identity, the randomly generated encryption keys themselves are securely deleted, and the hard drives in the machines on which all of these operations took place eventually physically destroyed. See http://www.asci.org/digital2001/sidal/sidal.htm.

operating system gets around to overwriting it. Indeed, because a file's physical storage location will change each time it is opened and modified, its earlier incarnations will also persist until such time as that data may be overwritten. The easiest way to recover data, therefore, is by simply locating a "deleted" file on the storage media after its entry has been stripped from the FAT but before any new data has been written to the same location. This is typically the way commercial recovery utilities work, hence the standard instruction to allow as little time as possible to elapse before attempting to restore a lost file. As hard drive capacities continue to increase, information will persist for longer and longer amounts of time on the surface of the platter before it is overwritten even once, thus expanding the window in which stored data remains recoverable.

But that alone does not account for the uniquely indelible nature of magnetic storage, or the uncompromising pronouncements of computer privacy experts like Michael Calonyiddes: "Electronic mail and computer records are far more permanent than any piece of paper."[46] Creating a file and saving it to a hard drive does not yield a simple one-to-one correspondence between the document (or file of whatever type) and its record on the disk. First, word processors and other productivity software routinely include an auto-save function that writes a snapshot of an open file to the disk at set intervals. The presence of such files is not always obvious: often they have opaque or arbitrary-seeming names (a copy of this document, for example, currently exists in one of my temporary directories as ~WRL0005.tmp). Nor do they always appear in standard directory listings (this same file would be invisible to me if the directory was configured to only display files with common extensions).[47] This phenomenon is sometimes known as "ambient data", the term emphasizing the way in which records accumulate on a file system absent the intervention of any single, conscious (human) agency. Most computers also use a portion of their hard disk as an extension of their RAM, a type of storage known as virtual memory or swap space. Forensic investigators recover all manner of otherwise-ephemeral matter, including passwords and encryption keys, from the swap space. So-called slack space—not to be confused with swap space—presents yet another opportunity for extracting remnants of supposedly

[46] Michael Caloyannides, *Computer Forensics and Privacy* (Norwood, MA: Artech House, 2001), 4.

[47] Caloyannides, 25.

long-discarded files. Data on a magnetic hard drive is stored in clusters of a fixed length; 4096 bytes is typical. (This is what accounts for the discrepancy between the actual size of a file and its "size on disk," as revealed by Windows Properties; even a one byte file—a single ASCII character—will require the allocation of a full 4096-byte cluster to store.) If a file is smaller than that (or larger, but not equivalent in size to some precise multiple of 4096), then the extra space in the cluster is filled out by information in the computer's RAM memory at the moment the file is committed to disk. But since files themselves are rarely the exact same size (and hence occupy variable numbers of clusters), it is also frequently possible to find the partial remains of earlier files at the end of a so-called cluster chain, a phenomenon sometimes known as "disk slack" (as opposed to file slack). A skilled investigator develops an instinct for where slack of either kind is to be found. The problem is exacerbated still further by the fact that in addition to temporary copies and other multiples of the actual file, metadata—the name of the file, the file type, date and time stamps—proliferates even more aggressively through the operating system, so even if the *content* of a file is completely erased it is still possible to recover evidence testifying to its past presence.[48] The interactions of modern productivity software and mature physical storage media such as a hard drive may finally resemble something like a quantum pinball machine, with a single simple input from the user sending files careening n-dimensionally through the internal mechanisms of the operating system, these files leaving persistent versions of themselves behind at every point they touch—like after-images that only gradually fade—and the persistent versions themselves creating versions that multiply in like manner through the system. There is, in short, no simple way to know how many instances of a single file are residing in how many states, in how many different locations, at any given moment in the operating system. Thus, as one textbook has it, "Deleted file information is like a fossil: a skeleton may be missing a bone here or there, but the fossil does not change until it is destroyed."[49] Nor is there any simple way to know how many metadata records of a file (or any of its ambient versions) exist. Given all this, it is not hard to see why one expert is left to conclude, "Secure file deletion on Windows platforms is a major exercise, and can only be part of a secure 'wipe' of one's entire hard disk. Anything less than that is likely to leave discoverable electronic evidence behind."[50]

Fungibility The preceding should make clear the extent to which a lay user's view of a file system—accessed only through standard directory structures or a Find function, or with the aid of menus, and manipulated using commands like Copy, Rename, and Remove—is optimized and impoverished, a partial and simplistic window onto the diverse electronic records that have accumulated on the surface of the magnetic disk. Because of this, when a hard disk is duplicated for forensic investigation it is not enough to simply copy the files in the usual manner (dragging and dropping the folders). Instead, an investigator will want to create a so-called bitstream image of the original file system. A bitstream is exactly that: every bit recorded on some original, physical instance of storage media is transferred in linear sequence to the copied image, whether it is part of a file currently allocated in the FAT or not. This means that all of the other ambient data on the original media is retained as part of the forensic object, including even (if the process is done right) data in "bad" or corrupted sectors no longer otherwise accessible. Since no forensic lab wants to work on the original source media and risk compromising its integrity, the proper execution of the imaging process is essential for creating legally admissible digital evidence. Legally, a bitstream copy of the original bits can usually stand in for the original digital object—what is known in courtroom parlance as documentary, rather than merely demonstrative, evidence.[51] While bitstream drive imaging and its legal status as documentary evidence would seem to reinforce the familiar postmodern argument about the digital simulacrum—copies without an original—it in fact underscores

[48] "[C]omputers," Caloyannides opines, "are a forensic investigator's dream because, in addition to the files themselves, they contain data about the data" (35). Caloyannides devotes particular attention to the "registry," which on Windows systems is actually a confederation of files that store basic, persistent information about the state of the user's system, including (potentially) indelible records of every piece of software installed (and removed), internet browsing histories, names and other personal identifying data, and so forth. A privacy activist as well as an author, Caloyannides makes it his business to describe in detail how to purge a registry of potentially incriminating content.

[49] Dan Farmer and Wietse Venema, *Forensic Discovery* (Upper Saddle River, NJ: Addison-Wesley, 2005), 159.

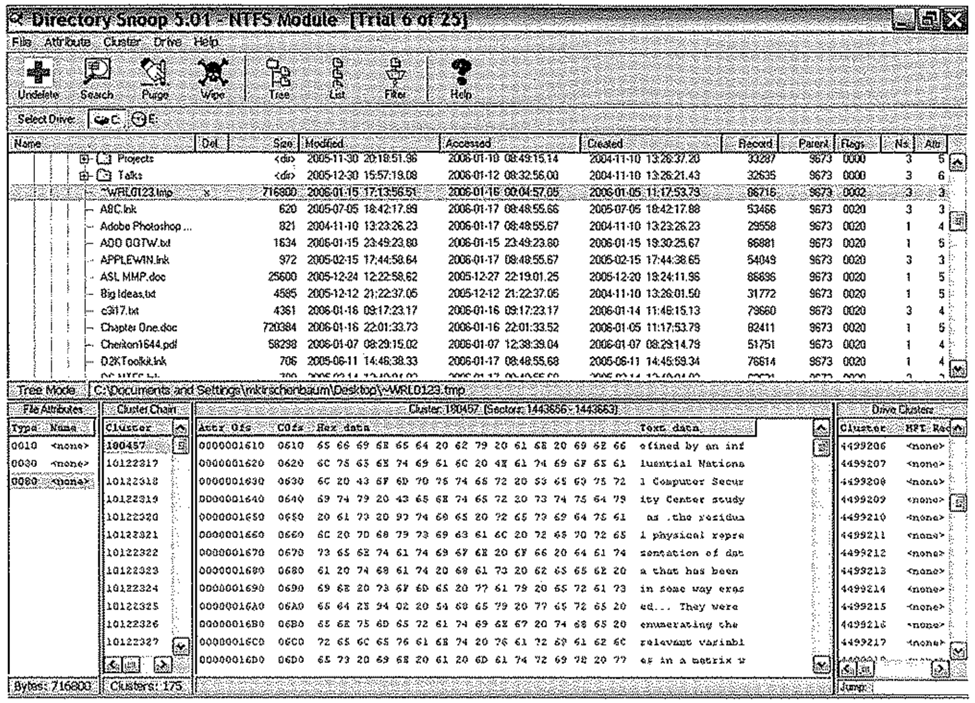
[50] Caloyannides, 28.

[51] See http://faculty.ncwc.edu/TOConnor/426/426lect06.htm for more on the distinction between documentary and demonstrative evidence in the context of applied computer forensics.

the heterogeneity of digital reproduction. The more mundane kinds of file reproduction we all perform in our daily interactions with a computer fall short of the forensic ideal.

Nor are these necessarily matters solely of relevance to legal investigation and standards of juridical evidence. In 2001, new media artist Joshua Davis (best known for praystation.com) released the Praystation Hardrive [*sic*], a limited·edition CD-ROM consisting of "original source files, art, text, accidents, epiphanies, all as they appear in Davis' own hard drive . . . 397 folders, 3637 files, 462 Meg of raw data: a snapshot of the studio of a major new media artist, frozen and time and delivered to you for exploration, study, inspiration, and for use in your own work."[52] Accompanied by liner notes pointing out highlights and packaged in a black plastic case vaguely reminiscent of a hard drive, the work is admirable in a number of ways: it manifests an open source ethos of creativity that feeds off of ever more capacious storage media, media that allows innumerable versions and layers and masters and derivatives to co-exist without the need to delete the extraneous matter to make room for more. It is also an invaluable historic document that captures a working set of software practices, the kind of artifact we need to learn to cultivate and appreciate. There is even a photo section featuring seemingly random personal photos of trips to San Diego and Paris. Rather than an executable launch platform, the only interface to the CD-ROM's content is the user's own desktop, where it is simply appended to the file system to be accessed via the normal directory navigation tools. Nonetheless, the data is not quite as raw as we are led to believe (figure 1.3). This is not a bitstream copy of the sort described above, and it is far indeed from a forensically sound copy of the hard drive itself, since none of the ambient data that would represent the systems-level workings of the files is present. Despite the conceit that we have been granted unmediated access to Davis's hard drive (or its digital surrogate), there is a greater artifactual distance between this copy and the original data objects than most users would commonly acknowledge. While I would not want to suggest that the failure to produce a bitstream simulacrum compromises Davis's project in any meaningful way, it does expose the merely rhetorical nature of some of its claims about the rawness of the data it provides. More importantly for my purposes, the example demonstrates the heterogeneity that I have asserted

[52] Text quoted from the Eastgate Systems catalog entry, http://www.eastgate.com/catalog/Praystation.html.



```
Name                     Del     Size  Modified                Accessed                Created                 Record  Parent  Flags  Ns  Attr
  ⊞ Projects                    <dir>  2005-11-30 20:18:51.36  2006-01-10 08:49:15.14  2004-11-10 13:26:37.20   33287    9673  0000    3   5
  ⊞ Talks                       <dir>  2005-12-30 15:57:19.08  2006-01-12 08:32:56.00  2004-11-10 13:26:21.43   32635    9673  0000    3   6
  ~WRL0123.tmp            x    716800  2006-01-15 17:13:56.51  2006-01-16 00:04:57.05  2006-01-05 11:17:53.79   86716    9673  0002    3   3
  ABC.lnk                         620  2005-07-05 18:42:17.89  2006-01-17 08:48:55.66  2005-07-05 18:42:17.88   53466    9673  0020    3   3
  Adobe Photoshop ...             821  2004-11-10 13:23:26.23  2006-01-17 08:48:55.67  2004-11-10 13:23:26.23   29558    9673  0020    1   4
  ADD OOTW.txt                   1634  2006-01-15 23:49:23.80  2006-01-15 23:49:23.80  2006-01-15 19:30:25.67   86881    9673  0020    1   5
  APPLEWIN.lnk                    972  2005-02-15 17:44:58.64  2006-01-17 08:48:55.67  2005-02-15 17:44:38.65   54049    9673  0020    3   3
  ASL MMP.doc                   25600  2005-12-24 12:22:58.62  2005-12-27 22:19:01.25  2005-12-20 18:24:11.96   86696    9673  0020    1   5
  Big Ideas.txt                  4585  2005-12-12 21:22:37.06  2005-12-12 21:22:37.06  2004-11-10 13:26:01.50   31772    9673  0020    1   5
  c3l17.txt                      4361  2006-01-16 09:17:23.17  2006-01-16 09:17:23.17  2006-01-14 11:46:15.13   79660    9673  0020    3   4
  Chapter One.doc              720384  2006-01-16 22:01:33.73  2006-01-16 22:01:33.52  2006-01-05 11:17:53.79   82411    9673  0020    1   5
  Cheriton1644.pdf              58298  2006-01-07 08:29:15.02  2006-01-07 12:38:39.04  2006-01-07 08:29:14.79   51751    9673  0020    1   4
  D2KToolkit.lnk                  706  2005-06-11 14:46:38.33  2006-01-17 08:48:55.68  2005-06-11 14:45:59.34   76614    9673  0020    1   4

C:\Documents and Settings\mkirschenbaum\Desktop\~WRL0123.tmp

Type  Name    Cluster   Cluster: 180457 (Sectors: 1443656 - 1443663)                                                Cluster  MFT Rec
                        Attr Ofs    COfs  Hex data                                         Text data
0010  <none>  180457    0000001610  0610  65 66 69 6E 65 64 20 62 79 20 61 6E 20 69 6E 66  efined by an inf    4499206  <none>
0030  <none>  10122317  0000001620  0620  6C 75 65 6E 74 69 61 6C 20 4E 61 74 69 6F 6E 61  luential Nationa    4499207  <none>
0080  <none>  10122318  0000001630  0630  6C 20 43 6F 6D 70 75 74 65 72 20 53 65 63 75 72  l Computer Secur    4499208  <none>
              10122319  0000001640  0640  69 74 79 20 43 65 6E 74 65 72 20 73 74 75 64 79  ity Center study    4499209  <none>
              10122320  0000001650  0650  20 61 73 20 97 74 68 65 20 72 65 73 69 64 75 61   as .the residua    4499210  <none>
              10122321  0000001660  0660  6C 20 70 68 79 73 69 63 61 6C 20 72 65 70 72 65  l physical repre    4499211  <none>
              10122322  0000001670  0670  73 65 6E 74 61 74 69 6F 6E 20 6F 66 20 64 61 74  sentation of dat    4499212  <none>
              10122323  0000001680  0680  61 20 74 68 61 74 20 68 61 73 20 62 65 65 6E 20  a that has been     4499213  <none>
              10122324  0000001690  0690  69 6E 20 73 6F 6D 65 20 77 61 79 20 65 72 61 73  in some way eras    4499214  <none>
              10122325  00000016A0  06A0  65 64 2E 2E 2E 20 54 68 65 79 20 77 65 72 65 20  ed... They were     4499215  <none>
              10122326  00000016B0  06B0  65 6E 75 6D 65 72 61 74 69 6E 67 20 74 68 65 20  enumerating the     4499216  <none>
              10122327  00000016C0  06C0  72 65 6C 65 76 61 6E 74 20 76 61 72 69 61 62 6C  relevant variabl    4499217  <none>
                        00000016D0  06D0  65 73 20 69 6E 20 61 20 6D 61 74 72 69 78 20 77  es in a matrix w

Bytes: 716800   Clusters: 175
```

Figure 1.3 View of the author's own desktop through a hex editor, here featuring a "deleted" "temporary" copy of the current file for this chapter. This kind of ambient data is not available on the Praystation CD, though it would have been present on Davis's file system. Screenshot by the author.

attends different kinds of digital copies, and serves to defamiliarize the Copy command itself—a command so mundane that we find it under every Edit menu, but which in fact condenses complex behaviors related to storage media, file integrity, and the various states data can inhabit.

The integrity of a bitstream image can be verified using as technique known as hashing. As we will see in the next chapter, hashing is a long-established electronic textual operation that has a role in domains ranging from error checking to cryptography. A hash algorithm generates a numeric value that is the mathematical surrogate for a particular bitstream. If the bitstream is altered in any way, by even so much as a keystroke, its hash value will be compromised, providing evidence of the tampering. Here is how one forensics textbook explains a cryptographic hash such as MD5 (developed by

Ronald Rivest, the MIT cryptographer who also worked on the popular RSA public key encryption standard):

> A cryptographic hash algorithm is a one-way form of encryption, taking a variable-length input and providing a fixed-length output. As long as the size of the original object is within the operational restraints of a particular implementation, it is statistically impossible for cryptographically secure hash algorithms to allow two different source files to have intersecting values, effectively making the hash value into the fingerprint of the original file. Such an algorithm is designed to be collision free, meaning that it is functionally impossible to create a document that has the same cryptographically secure hash value as another document. Because of this characteristic, a hash value can serve as a surrogate for the object it is derived from.[53]

Hashing thus demonstrates that electronic objects can be algorithmically individualized. If Locard's Exchange Principle is the first basic tenet of forensic science, the second is individualization—the core faith that no two objects in nature are ever exactly alike, and no two objects ever break or wear down in the same way. Individualization, which we will explore in more detail later, is the principle underlying standard identification techniques like fingerprinting and DNA. An MD5 hash, which yields 2^{160} different values, is in fact a more reliable index of individualization than DNA testing.[54]

Fixity and Fluidity We tend to fixate on the fluidity of electronic text, its malleability and putative instability. Yet there is nothing essentially fluid about data in an electronic environment, and there are important areas in which the stability and verifiability of electronic documents is essential, ranging from secure e-commerce to security and defense. Anyone who has ever needed to edit a file in a directory to which they do not have access rights knows exactly how stubbornly resilient electronic text can suddenly become. On a more specialized level, Intelligent Computer Solutions has developed a product line popular in the forensics and law enforcement communities called Image MaSSter DriveLocks:

> Designed exclusively for Forensic applications, the DriveLock IDE device provides a secure hardware write protect solution for hard disk drives. Sensitive Forensic hard disk drive data can be previewed, acquired or analyzed without the possibility of altering the drive's contents. The device is designed to block write commands sent to hard disk drives connected through the computer's P-ATA interface. No special software is needed. The unit is compact and is easily portable.[55]

The existence of tools and technologies like DriveLocks and digital watermarking or signatures should remind us that the conditions governing electronic textuality are formal conditions—artificial arrays of possibility put into play by particular software systems. Just as the electronic textual field can be thrown open to revision by virtue of its susceptibility to formal manipulation, so too can this potential be—formally—foreclosed. (The ultimate arbitrariness of the situation is perhaps best brought home by the inevitable existence of the Image MaSSter's antiproduct, WipeMaSSter, marketed by the same parent company and "designed as a compact, standalone hardware solution for sanitizing and erasing drive data for up to nine drives simultaneously at speeds exceeding 3GB/min.")[56]

Secure digital document design—creating documents that can be electronically signed and sealed as guarantors of authenticity—is currently a thriving field. And questions of authenticity are directly related to an electronic object's ability to not only resist but also to expose tampering. Johanna Drucker, in an illustrated essay in the type design magazine *Émigré*, makes the point with characteristic acuity: "The authority of written documents . . . does not depend upon their pristine and unaltered condition. Quite the contrary—it is the capacity of the material documents to record change that makes them such believable witnesses."[57] Michael Hancher, in a perceptive essay that amply demonstrates the fruits of a textual scholar's encounter with electronic media, argues much the same point, by way of the eighteenth century legal theorist William Blackstone.[58] For Blackstone, the

[53] Warren G. Kruse II and Jay G. Heiser, *Computer Forensics* (Boston: Addison-Wesley, 2002), 89; emphasis in original.

[54] Kruse and Heiser, 89.

[55] http://www.atp-p51.com/html/drivelock.htm.

[56] http://www.atp-p51.com/html/wipemasster.htm.

[57] Johanna Drucker, "The Future of Writing in Terms of its Past: The New Fungibility Factor." *Émigré* 35 (Summer 1995).

[58] Michael Hancher, "*Littera Scripta Manet*: Blackstone and Electronic Text," *Studies in Bibliography* 54 (2001): 115–132.

most suitable surface for deeds, contracts, and other official documents was paper rather than stone, leather, wood, or other writing surfaces, which were not unknown in an era when linen rag paper was expensive and comparatively scarce. Paper, however, has the virtue of being both durable and secure—secure meaning precisely that it is fragile and vulnerable enough to readily reveal tampering. Stone, on the other hand, is inviolate—stains could be lifted or otherwise expunged without physically damaging the underlying surface, thus affording no guarantee that the writing has not been tampered with (122–123). Hancher then extends the argument to electronic text and suggests neither that it is not durable (he realizes that durability consists in multiplication as well as persistence), nor that it is not secure (he is aware of state-of-the-art technologies using electronic keys and signatures). Rather, his point is that authenticating electronic documents requires the services of an expert "postmodern technician" because "it deals with a disembodied reality inaccessible to and unassessable by the laity" (130). The fact that one can educate oneself in the particulars of electronic keys and signatures notwithstanding, most users will simply not have entrée to the mechanisms governing their individual interactions with supposedly secure and authentic electronic information. According to Hancher, one must ultimately take the security of electronic documents on "faith" (131).

Regardless, electronic document security—situated at the intersection of encryption, computer security, e-commerce, digital rights management, and digital archives and records management—can only underscore the limited and arbitrary nature of any medial ideology that celebrates only the fluidity and fungibility of electronic text. Powerful and well-financed constituencies are lobbying for a very different electronic textual condition, and the research and development is well under way.

None of what I have been describing yet accounts for the physical properties of magnetic media, or for the behavior of the drive's actual writing mechanism (which is responsible for a phenomenon known as shadow data, when bit representations turn out to be imperfectly overwritten). I will discuss these in the next section.

Inscription and Instrumentation

A document such as the Clearing and Sanitization Matrix is born of electronic data's eventual (and often immediate) status as inscribed trace; that is, an intervention in or modification of a physical substratum. We see this vividly in the phenomenon of data remanence, the physical remains of data deposited on computer storage media:

> As early as 1960 the problem caused by the retentive properties of AIS [automated information systems, i.e., computers] storage media (i.e., data remanence) was recognized. It was known that without the application of data removal procedures, inadvertent disclosure of sensitive information was possible should the storage media be released into an uncontrolled environment. Degaussing, overwriting, data encryption, and media destruction are some of the methods that have been employed to safeguard against disclosure of sensitive information. Over a period of time, certain practices have been accepted for the clearing and purging of AIS storage media.[59]

This concern is not limited to national security and shadowy government agencies. Corporations and institutions of all kinds, not to mention private individuals, routinely discard computers, often with little or no attention to the risk of data remaining intact on the hard drive. One recent study examined a sampling of discarded hard drives and found that nearly all of them contained sensitive information that was recoverable to varying degrees.[60] Indeed, sometimes not even a token attempt had been made to delete files from the disk. In other cases the utilities used were insufficient to address the ambient data that had accumulated. Contrary to popular belief, initializing a disk does *not* erase or overwrite all of its data; it only erases the FAT and resets the basic formatting information. Actual data remains on the disk, and can be recovered using well-known techniques.[61]

[59] National Computer Security Center, "A Guide to Understanding Data Remanence in Automated Information Systems," NCSC-TG-025, http://all.net/books/standards/remnants/index.html.

[60] Simson L. Garfinkel and Abhi Shelat, "Remembrance of Data Passed: A Study of Disk Sanitization Practices," *IEEE Security and Privacy* (January–February 2003): 17–27. The authors conclude: "With several months of work and relatively little financial expenditure we were able to retrieve thousands of credit card numbers and extraordinarily personal information on many individuals. . . . If sanitization practices are not significantly improved, it's only a matter of time before the confidential information on repurposed hard drives is exploited by individuals and organizations that would do us harm" (27).

[61] See Garfinkel and Shelat, 17–27.

The considerations in play here go a step beyond those discussed in the previous section, where the problem of deleting data was strictly a function of its proliferation under a swarm of system processes not directly controlled by the user. Data remanence is also a function of the physical properties of storage media and the difficulty of reversing or obscuring what are tangible interventions in a physical medium. "Virtually all erasures can be detected by a thorough examination," wrote Ordway Hilton in his *Scientific Examination of Questioned Documents* (96). But he may as well have been talking about computer storage media. "You can't really erase a hard drive," unequivocally state the authors of one computer forensics textbook, likening it to the way a child's Etch A Sketch retains the images of previous drawings.[62] In fact you *can* erase a hard drive, but it is a deliberate and painstaking process, best attempted with the proper tools by an expert who understands the full extent of the issues involved. (Better still, perhaps, to simply run it over with a tank, as the NSA originally suggested, though modern data recovery abounds with seemingly miraculous stories of data extracted from hard drives subjected to near-Biblical levels of fire, flood, and blunt force trauma.[63] The recoveries performed on the World Trade Center hard drives are one such example.)

The paper that made much of this common knowledge within computer security and privacy circles was published in 1996 by Peter Gutmann, a researcher at the University of Auckland. Titled "Secure Deletion of Data from Magnetic and Solid-State Memory," Guttman's paper begins by making

[62] Kruse and Heiser, 77.

[63] Some of the best tales of data recovery I know come from the Web site of a company called DriveSavers, which includes testimonials from the likes of Keith Richards, Sean Connery, Sarah Jessica Parker, Sting, Industrial Light and Magic, and Isaac Hayes. Yet one client in particular stands out. As the Web site tells it, a technician was working on the hard drive of someone he assumed must be a hardcore *Simpsons* aficionado: the disk was full of character stills, icons, animations, renderings, etc. Then the technician came across a folder labeled "Scripts." It turned out that the drive belonged to a writer for the show, and the damaged disk contained the only copies of the scripts for twelve then unproduced episodes. The scripts, which included the famous season finale "Who Shot Mr. Burns," were all recovered successfully. Thus at least one signature artifact of pop culture owes its existence to the art of forensic data recovery. Other entries on the site tell of recovering data from laptops submerged in the Amazon river, scorched in house fires, and overrun by an 18-wheel truck. See http://www.drivesavers.com/fame/index.html.

the point that while most of what we know about data remanence comes from intelligence agencies, it is not in these sources' best interests to disclose everything they actually know.[64] Therefore he cautions against underestimating official capabilities. His point of departure is an esoteric technology known as magnetic force microscopy, or MFM. Pioneered in the late 1980s, magnetic force microscopy was and is the method of choice for imaging data representations recorded on magnetic media. Its primary application is not forensic recovery but industrial research and development: MFM studies are an integral part of evaluating laboratory advances in magnetic recording. MFM is actually an umbrella term for several closely related procedures and technologies, all based on the scanning tunneling microscope (STM; a variety of electron microscope), and it in turn offers only one of several known methods for imaging magnetic phenomena. A magnetic force microscope, as the name implies, is essentially a feedback device. A flexible probe, made of iron, is positioned just above the surface of a magnetic media. Figure 1.4 is an example of the kind of output generated.

What we see here are not simply bits, but patterns of magnetic flux reversals, a number of which may be necessary to constitute a single bit (which I discuss in more detail in chapter 2). Thus while bits are the smallest symbolic units of computation, they are not the smallest inscribed unit, a disjunction that underscores the need to distinguish between the forensic and the formal in discussions of computational materiality.

In order to generate these images, the tip of the microscope is moved above the surface of the platter, typically at a distance of only a few nanometers. Electrons "tunnel" from the surface of the platter to the tip of the probe, repelling or attracting it; the microscope, meanwhile, exerts greater or lesser force to keep the tip at a constant distance from the surface. Thus, the energy

[64] Peter Gutmann, "Secure Deletion of Data from Magnetic and Solid-State Memory," first published in the *Sixth USENIX Security Symposium Proceedings*, San Jose, California, July 22–25, 1996. Online at http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html. Gutmann's paper has not gone without challenge and critique. Daniel Feenberg, a researcher at the National Bureau of Economic Research, contends that it is "overwrought," pointing out that there has never been a known, actual instance of MFM technology being employed for forensic data recovery by an intelligence agency or other government entity. See http://www.nber.org/sys-admin/overwritten-data-guttman.html. Nonetheless, MFM serves to demonstrate the irreducibly physical basis of digital inscription.

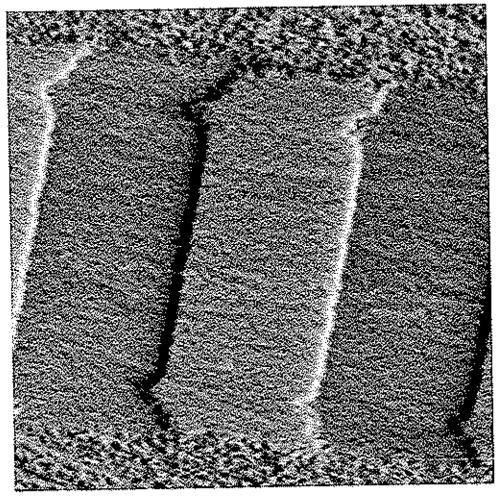


Figure 1.4 Magnetic force microscopy (MFM) image of bits on the surface of a hard disk. 15 × 15 microns (μm). http://web.archive.org/web/*http://www.boulder.nist.gov/magtech/mfm.htm, Mar 06 2005.

the probe expends is the basis for measuring the magnetic force. The raw images are then subjected to layers of higher-level image processing to generate the kind of image depicted here. As a visual rendition of the magnetic fields active on the surface of the source media, an MFM image is ultimately more akin to a physician's ultrasound than the detective's magnifying glass. The bits themselves prove strikingly autographic, all of them similar but no two exactly alike, each displaying idiosyncrasies and imperfections—in much the same way that conventional letterforms, both typed and handwritten, assume their own individual personality under extreme magnification. It seems counterintuitive to think of bits as revealed artifacts—rectangular, with an aspect ration of 8:1, measuring here about 4 × .5 microns—yet this is where Hayles's "flexible chain of markers bound together by the arbitrary relations specified by the relevant codes" comes to its end, as an actual inscribed trace. (The smallest possible bits we can write are probably about 10 nanometers, 400 times smaller than what is the norm today. At this level bits will approach what scientists called the superparamagnetic limit—the point at which a physical surface area is no longer capable of retaining a magnetic charge.)

While MFM sounds like an exotic technology, Gutmann suggests barriers to its use are less than one might imagine:

> Even for a relatively inexperienced user the time to start getting images of the data on a drive platter is about 5 minutes. To start getting useful images of a particular track requires more than a passing knowledge of disk formats, but these are well-documented, and once the correct location on the platter is found a single image would take approximately 2–10 minutes . . . (2)

Guttman concludes that "Faced with technologies such as MFM, truly deleting data from magnetic media is very difficult" (2). Data remanence of the sort that MFM exploits is ultimately a function of the physical properties of the magnetic substrate and the material limitations of the drive's write technology—the "inability of the writing device to write in exactly the same location each time"—as well as variations in sensitivity and the strength of magnetic fields on the source media (2). This effect satisfies the forensic principle of individualization, which insists upon the absolute uniqueness of all physical objects. The core precepts of individualization construct a hard materiality of the kind that ought to resonate with textual scholars and others in the traditional humanities: "No two things that happen by chance ever happen in exactly the same way; No two things are ever constructed or manufactured in exactly the same way; No two things ever wear in exactly the same way; No two things ever break in exactly the same way."[65] That the scale here is measured in mere microns does not change the fact that data recording in magnetic media is finally and fundamentally a forensically individualized process.[66]

[65] Harold Tuthill *Individualization: Principles and Procedures in Criminalistics* (Salem, OR: Lightning Powder Co., 1994), 20.

[66] On November 14, 2003, I visited Professor Romel Gomez at the Universiy of Maryland's Laboratory for Applied Physics. Professor Gomez heads the Lab's nanomagnetics group. I observed a Digital Instruments (now Veeco) magnetic force microscope in action. The device was recognizable as a microscope, with familiar elements such as a stage and ocular tubes. Three monitors provide views: one shows an optical magnification of the surface of the media sample, the second displays feedback from the instrumentation and settings, and the third displays reconstructed images of the magnetic data, both AFM and MFM. The process is time- and labor-intensive, more so than Guttman seems to suggest: acquisition rates hover around 1 bit (not byte) per second, and the surface area of a sample is small—perhaps five square millimeters.

The phenomenon that speaks most directly to electronic data's status as an individualized inscription is well-documented in the MFM literature: tracking misregistration. It occurs in two different forms. Large misregistrations leave the remnants of earlier data representations plainly visible along the edges of the track, exposed to forensic detection and recovery—a classic palimpsest effect, sometimes known as shadow data. "Given intrinsic limitations in the positioning of the head, this effect might be more ubiquitous than previously realized."[67] Thus when Bruce Clarke, a sophisticated theorist, writes "material . . . if deleted and overwritten, leaves no scratch on any surface" (31) he is correct only in the narrow, literal sense that electronic data does not impinge on the surface of its substrate in the form of a scratch.[68] In addition to the presence of shadow data, when new bits are recorded the positioning of the write head may also be off just enough that the magnetic field is strong enough to erase the old data, but not strong enough to successfully record the new data. This creates what is known as an erase band along the edges of the data track, a magnetic anomaly that has a characteristic signature when viewed with MFM imaging (see figure 1.5). The erase band is the blurred area near the top of each image where there is no definite magnetization of the source media. Also visible here are data imperfectly overwritten due to larger (relatively speaking) alterations in the positioning of the head as it makes successive passes over the same data track.

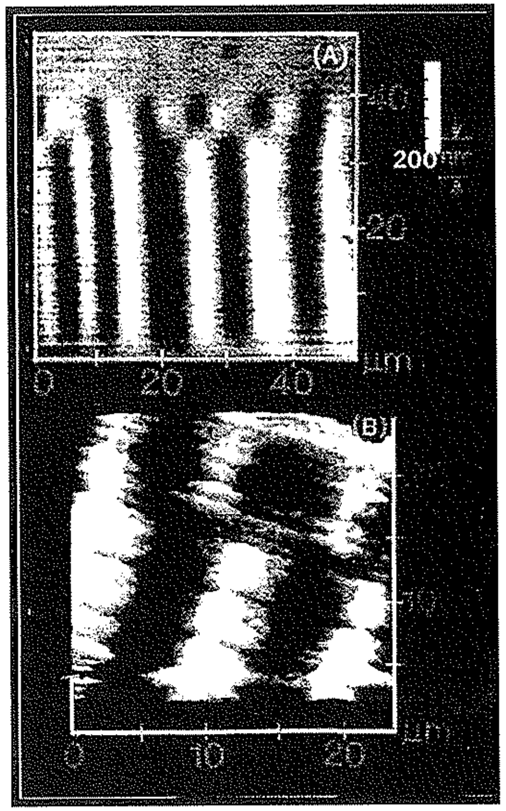


Figure 1.5 MFM erase band study. Taken from R. D. Gomez, E. R. Burke, A. A. Adly, and I. D. Mayergoyz. "Magnetic Force Scanning Tunneling Microscope Imaging of Overwritten Data," *IEEE Transactions on Magnetics* 28 (September 1992): 3141.

The conclusion researchers have reached describes a condition serving to distinguish magnetic recording from other kinds of inscription, such as ink staining a fibrous sheet of paper or the grooves incised on a wax cylinder:

> For small tracking misregistrations, localized remnants of overwritten data may no longer be distinctly detectable but continue to perturb the system by its influence on the characteristic trackwidth variations of the newly created data. Thus, it is quite

If we do the math—eight bits in a byte—we can see that we might, assuming optimal conditions, be able to image seven or eight bytes per minute. A single 512 byte sector would require well over an hour to image completely. A relatively modest ten kilobyte text file would require 24 hours of continuous imaging under optimal conditions. A 1 MB media file would take months. Though recoveries of complete files are theoretically possible (through what is known in the trade as "heroic efforts"), the process would be extraordinarily painstaking and take weeks or months. For a good general introduction to MFM, see Gomez et al., "Magnetic Force Scanning Tunneling Microscopy: Theory and Experiment," *IEEE Transactions on Magnetics* 29 (November 1993): 2494–2499. A new technology, known as spin-stand imaging and capable of attaining much higher acquisition rates, is in development. See I. D. Mayergoyz et al., "Spin-Stand Imaging of Overwritten Data and its Comparison with Magnetic Force Microscopy," *Journal of Applied Physics* 89 (June 2001): 6772–6774.

[67] Gomez et al. (1993), 2499.

[68] Bruce Clarke, "Thermodynamics to Virtual Reality" in Bruce Clarke and Linda Dalrymple Henderson, eds., *From Energy to Information: Representation in Science and Technology, Art, and Literature*, (Stanford: Stanford University Press, 2002): 17–33.

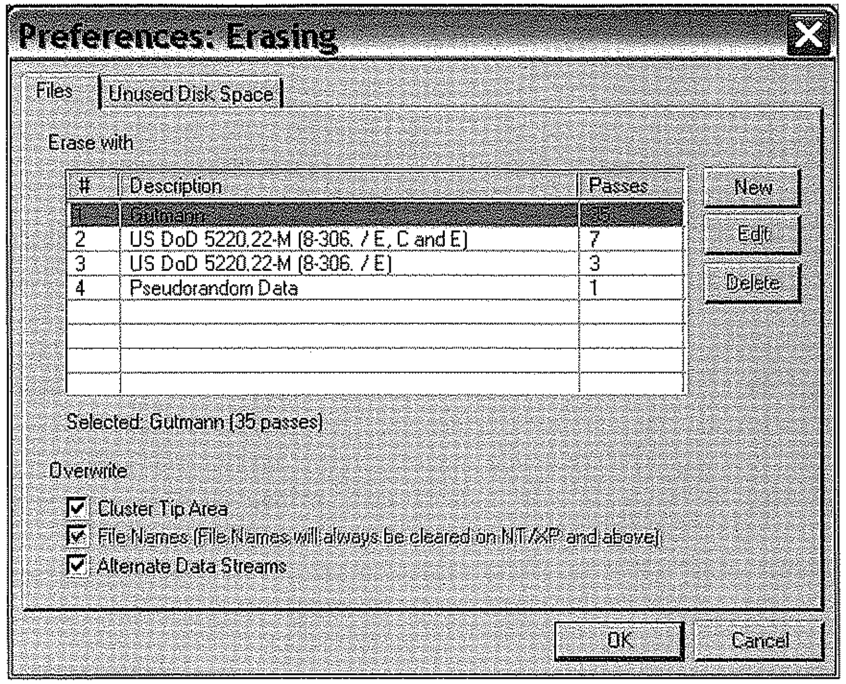
possible that even with direct overwrite...complete elimination of the effects of previous data may not be achieved.[69]

Gutmann puts it this way: "Each track contains an image of everything ever written to it, but the contribution from each 'layer' gets progressively smaller the further back it was made" (3). In other words, magnetic inscription is a temporal as well as a planographic intervention, whereby even data that has been overwritten continues to resonate as a result of the ongoing oscillation of the magnetic field. This basic property of magnetic media is known as hysteresis. Gutmann's solution involves not erasing, but writing and rewriting—thus repressing the lingering effects of earlier data. The bulk of his paper develops a set of 35 patterns designed to ensure that ones are overwritten with zeros overwritten with ones, while zeros are overwritten with ones overwritten with zeros. This goes on through so many layers of recursion that eventually the ability of a scanning device (like MFM) to detect significant enough fluctuations in field strength to recuperate earlier data patterns is negated. Gutmann's patterns have since become canonical, so much so that disk sanitizing utilities encode them as an explicit option, as is visible in figure 1.6.

In many respects, MFM represents the continuation of a scientific imaging tradition dating back to Faraday's drawings of lines of magnetic force in the 1830s. Digital inscription is itself inseparable from practices of instrumentation, and the history of science and technology is marked by continuous attempts to visualize and render such insubstantial phenomena as the ether, electricity, and electromagnetism. Indeed, the ether into which digital objects are often said to vanish is a historically constructed and contested site, with a rich tradition of visualization and imaging/imagining that erupted in the late nineteenth century.[70] One outcome of an encounter with a technology like MFM is that "the virtual" turns out to be a more heterogeneous category than we may have first thought, since at least some of what is usually subsumed in that category is in fact not virtual, but only very small, so small as to be invisible to wave length optics even under the most extreme magnification.[71]

[69] Gomez et al. "Microscopic Investigations of Overwritten Data," *Journal of Applied Physics* 73.10 (May 1993): 6001–6003.

[70] See, for example, Bruce J. Hunt, "Lines of Force, Swirls of Ether" in Bruce Clarke and Linda Dalrymple Henderson, eds., *From Energy to Information: Representation in Science and Technology, Art, and Literature*, (Stanford: Stanford University Press, 2002): 99–113.



Erase with

#	Description	Passes
1	Gutmann	35
2	US DoD 5220.22-M (8-306. / E, C and E)	7
3	US DoD 5220.22-M (8-306. / E)	3
4	Pseudorandom Data	1

Selected: Gutmann (35 passes)

Overwrite

☑ Cluster Tip Area
☑ File Names (File Names will always be cleared on NT/XP and above)
☑ Alternate Data Streams

Figure 1.6 Preferences window of a popular utility for secure deletion called Eraser. The options for the 35 Gutmann patterns are available, as are various DoD-sanctioned patterns. Note too the option to erase unused disk space ("slack") alongside of actual files, something that is impossible with Windows systems commands alone. Screenshot by the author.

[71] The history of microscopy in the forensic field of questioned document examination is instructive here. Recent textbooks treat the microscope briefly and unremarkably, assuming the investigator's knowledge of and readiness to employ the instrument. For example, David Ellen's *The Scientific Examination of Questioned Documents: Methods and Techniques* (New York: Wiley and Sons, 1989) matter-of-factly states: "Magnification is an important part of document examination. Some enlargement can be produced by the use of photography and video methods, and also by a magnifying glass, but the optical microscope is the most used tool of the examiner" (163). Hilton, meanwhile, devotes none of his pages to the microscope as such, but simply furnishes examples of its application throughout the text. This was not always the case, however. Albert S. Osborn's *Questioned Documents*, whose first edition was published in 1910, gives over an entire chapter to the role of the microscope, misconceptions about it, and the dangers of its potential

Yet magnetic force microscopy is also unabashedly an instrument in Bruno Latour's sense of the word: "I will call an instrument (or inscription device) any set-up, no matter what its size, nature and cost, that provides a visual display of any sort in a scientific text."[72] For Latour, it is essential that the instrument assumes a rhetorical disposition as evidence marshaled in the service of scientific discourse. This is the forensic function of the MFM images, not only in courts of law, but in science and engineering circles, where they contribute to insights related to the low-level physical properties of magnetic recording. Entire articles in the literature are given over to improving the computer-mediated image processing techniques used to render the images resulting from the MFM process, which it must be remembered are not optical magnifications but force-feedback renderings.[73] Ultimately, as Nathan Brown presciently argues, the scanning tunneling microscope (of which MFM is a subclass) "constitutes an event in the history of writing machines insofar as it makes marks on a scale *beyond* optics, at which visual (re)presentations are predicated on the radical priority of a haptic interface" (175; emphasis in original).[74] Our most persuasive evidence for the autographic individualization of bit-level digital inscription comes not from sight, but from the instrumental touch of the mechanism.

abuse. The following prose is characteristic: "The objections to the use of the instrument usually are based upon the somewhat natural but erroneous idea that if a thing exists that is really significant it can be seen by unaided vision. It seems to be overlooked by those who object to the microscope that ordinary spectacles are simply lenses placed between the eye and the object looked at . . . and that the most elaborate and complicated microscope is nothing more than an extension of this same principle. To be consistent one who objects to the use of the microscope should also insist that the judge and jury should be compelled to remove spectacles before examining a document that is questioned in a court of law" (74). Thus Osborn wants to buttress the spectrum of the "physical" so that the evidence refracted through the compound lenses of the microscope is finally no less persuasive than what might be seen with the naked eye, or the quotidian intervention of spectacles.

[72] Latour, *Science in Action*, 68.

[73] For example, I. D. Mayergoyz, A. A. Adley, and R. D. Gomez, "Magnetization Image Reconstruction from Magnetic Force Scanning Tunneling Microscopy Images," *Journal of Applied Physics* 73, no. 10 (May 1993): 5799–5801.

[74] Nathan Brown, "Needle on the Real: Technoscience and Poetry at the Limits of Fabrication." *Nanoculture*, ed. N. Katherine Hayles (Bristol, UK: Intellect Books, 2004): 173–190.

Coda: CTRL-D, CTRL-Z

The Consumer Electronics Show, Las Vegas, Nevada, 1978. The most popular personal computers on the market are Radio Shack's TRS-80, the Commodore PET, and the Apple II. All use off-the-rack television sets for their displays and a standard cassette recorder for data storage. After driving all day, Apple's Steve Wozniak and Randy Wigginton arrive at the convention hall. The centerpiece of their booth is the Disk II, the first floppy disk drive for the home computer market. In time, the Disk II would become as important and iconic a part of the Apple II's identity as the computer itself. Woz and Wigginton, working straight through the previous week, have written some wildly inventive routines that dramatically improve the response times of the high-end corporate disk systems the drive is modeled on, notably replacing hard sectoring—which keyed data storage geometries to a hole physically punched into the disk—with so-called "soft sectoring," which allowed DOS to arrange the physical media however it saw fit. Apple historian Steven Weyrich chronicles what happened next:

> When they got to Las Vegas they helped to set up the booth, and then returned to working on the disk drive. They stayed up all night, and by six in the morning they had a functioning demonstration disk. Randy suggested making a copy of the disk, so they would have a backup if something went wrong. They copied the disk, track by track. When they were done, they found that they had copied the blank disk on top of their working demo! By 7:30 am they had recovered the lost information and went on to display the new disk drive at the show.[75]

Thus the disk handling routines that took the nascent personal computer industry by storm were accidentally overwritten on the very morning of their public debut—but recovered and restored again almost as quickly by those who had intimate knowledge of the disk's low-level formatting and geometry. Nowadays we toggle the CTRL-D and CTRL-Z shortcuts, deleting content and undoing the act at a whim. Gone and then back again, the keyboard-chorded Fort and Da of contemporary knowledge work.

[75] http://apple2history.org/history/ah05.html.

The perceived volatility of electronically recorded data is one of the signature affordances of new media, from casual users who "lose" files when hard drives crash or network accounts expire, to librarians, archivists, and others charged with the preservation of cultural heritage. Computer forensics counteracts this anxiety, teaching investigators to evaluate the relative stability or vulnerability of data in memory states and storage locales throughout an operating system. This tension between erasable and indefinitely retainable data storage was explicitly anticipated as a key element of designing usable digital computing machinery by pioneers like Norbert Wiener.[76] As we will see in the next chapter, the hard disk drive was a landmark achievement in computer engineering precisely because it offered a solution for erasable but nonvolatile random access storage. Computing is thus situated within a millennia-long tradition of reusable writing technologies, a tradition which also includes wax writing tables, graphite pencils, and correctible typewriter ribbons. We see this in the names of popular file removal utilities: Wiper, FileWiper, BCWipe, CyberScrub, Eraser, Shredder, Shred-It, Shred-X, Burn. Peter Gutmann's patterns for overwriting magnetic media are not so different in their way from the recipe for the ink and varnish used to create erasable surfaces for Renaissance writing tablets.[77]

For our purposes, contrasting experiences revolving around the erasure and the restoration of digital data can be usefully parsed as differences of forensic and formal materiality. Whereas formal materiality depends upon the use of the machine's symbolic regimen to model particular properties or behaviors of documents or other electronic objects (CTRL-Z thereby allowing one to "undo" a deletion by recapturing an earlier, still-saved state of file), forensic materiality rests upon the instrumental mark or trace, the inscription that is as fundamental to new media as it is to other impositions and interventions in the long technological history of writing. The Clearing and Sanitization Matrix with which we began the chapter clearly establishes the heterogeneity of digital inscription—not just in the range of potential source media, but in the variety of procedures used to create, destroy, and recover data. Overwriting information is not the same as degaussing it. Degaussing methods employ two basic kinds of magnetic fields, alternating (AC) and unidirectional (DC); the strength of the field required to reset the magnetic media to its unrecorded state is known as the media's coercivity, measured in oersteds (Oe). A weak degausser cannot securely erase media rated at a coercivity above 750 Oe, as most hard disks are. Likewise, overwriting may use a variety of different patterns and schemas to ensure that every bit is superimposed by its symbolic inverse. The debate over the effectiveness of these different patterns, and the effort to develop new, even more effective ones, is ongoing. The point is one that will be familiar to any student of writing technologies: writing practices engender an eruption of tools and techniques to fix, expunge, and recover their meaning-bearing marks and traces.

Formally then, electronic data is pernicious by virtue of its susceptibility to symbolic propagation in an environment explicitly built and engineered to model ideal conditions of immateriality (the essence of digital computing, as we will discuss in chapter 3). Forensically, electronic data is survivable by virtue of both dramatically expanding storage volumes (which make it trivial to retain redundant copies of an electronic object) and the limits of the material mechanism, as revealed in the spectral erase bands visible in the MFM images. Here, at tolerances measured in microns, is where we begin to locate the forensic imagination.

[76] Wiener's exact prescription from *Cybernetics* (1964) was as follows: "That the machine contain an apparatus for the storage of data which should record them quickly, hold them firmly until erasure, read them quickly, erase them quickly, and then be available immediately for the storage of new material" (4).

[77] For the key study of erasable writing tablets, see Peter Stallybrass, Rogier Chartier, John Mowery, and Heather Wolfe, "Hamlet's Tables and the Technologies of Writing in Renaissance England," *Shakespeare Quarterly* 55, no. 4 (Winter 2004): 379–419.