3 *New Media Bodies*

Card Stock

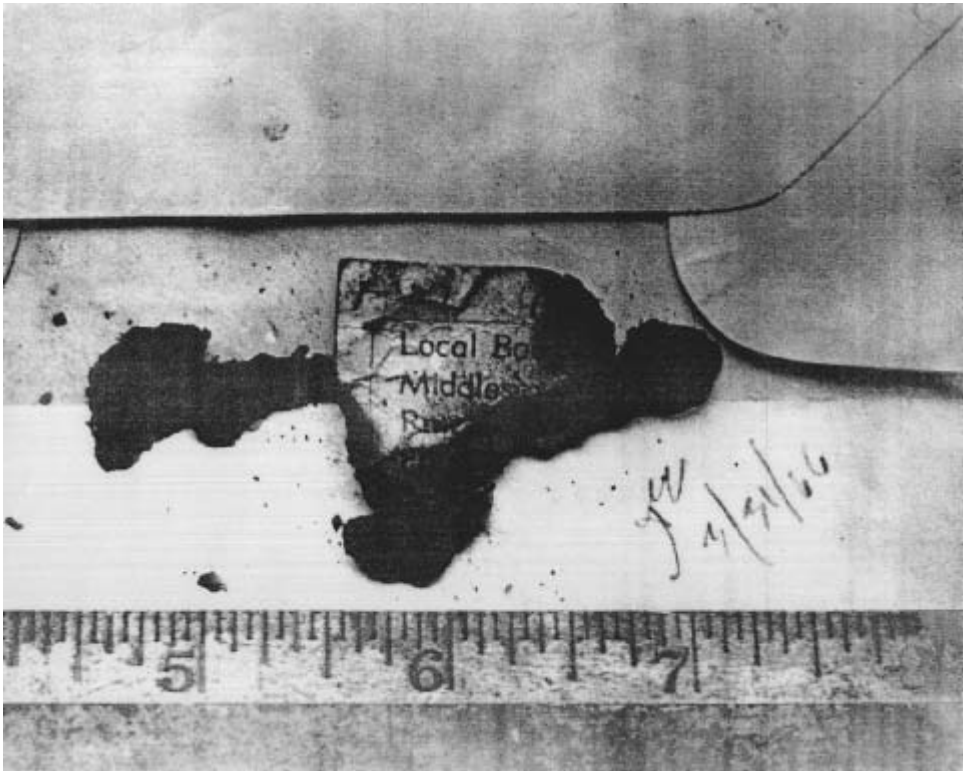
In its 1968 decision *United States v. O'Brien,* the U.S. Supreme Court tangled with free speech as a question of bibliography. The case concerned a young man in Boston who had publicly burned his draft card, and the Court was asked to decide whether his act might not be constitutionally protected under the First Amendment. The Earl Warren Court ruled against David O'Brien seven to one, and in the course of his decision, the chief justice chided the man for having been "unrealistic" in his characterization of draft cards as "so many pieces of paper" intended to notify draftees of their registration and then "retained or tossed in the wastebasket according to the convenience or taste of the registrant." These were not ephemeral notifications, not mere messages; they were lasting and indelible certificates, made so by a legitimate act of Congress. The point of the Selective Service card, the Court reasoned, was not just what the card said but also, more particularly, *that* it said: the physical integrity of this small white card made a difference because only as a physical (that is, bibliographic) body could it reliably facilitate "the smooth functioning of the system" of which it was part. That system, the Selective Service System (SSS), matched bibliographic bodies one-to-one with eighteen-year-old male human bodies in a highly rationalized way.

The Court's reasoning was tortured in several respects. On First Amendment grounds, the line the Court drew between speech and nonspeech has proved impossible to maintain in subsequent rulings on related cases.[1] On bibliographic grounds, the Court pointed to a number of different supposed functions that "would be defeated" by the destruction or mutilation of the cards. If draft cards were destroyed, they could not serve "as proof," Warren maintained, in a way that was "easy and painless" for draftees as well as "just and efficient" from the perspective of the SSS. As such, the cards worked like "receipts" that

Figures 3.1a and 3.1b *United States v. O'Brien* exhibits 1a and 2a: draft-card burning and the "charred remains," photographed by the FBI, March 31, 1966. (*Source:* Library of Congress.)

could certify a draftee's status in a "rapid and uncomplicated" manner by keeping necessary information handy. Further, if the cards were destroyed, they could not offer the "continual reminders" that they did, alerting registrants that they should be in touch with their draft boards about changes in status or address. And finally, if the cards were destroyed, it would "obviously" be much more difficult to identify abuses like alteration or forgery, which were clearly against the law.

In calling the Court's interest "bibliographic," I mean to underscore the complexity of its investment in the draft cards as meaningful, paper-based, textual forms. As texts, draft cards seem to have inhabited a vast and murky middle ground between two poles: the idea



of a pure text, on the one hand, and some sort of nontext, on the other. O'Brien had opted for the pure-text extreme, according to Chief Justice Warren, when he argued that the cards were only meaningful for—only meaningful *as*—the information they contained. The opposite extreme would be an untext of some kind, like a bookmark, which contains no information itself, but functions *in its body* as a meaningful instrument.[2] At one pole there would be only meaning—unadulterated information, content, message, data—and at the other there would be only matter—an empty piece of paper, a blank, performing its function. Everything in between the two extremes is sign plus supplement, meaning and/as material. (That ugly "and/as" is necessary because meaning and materiality are mutual and not distinct.) Some of the murkiness of the middle ground between the two extremes is suggested in part by the Court's weird elaborations, such as the implication that destroying the cards might abet forgers of them. Its vastness is indicated by the presence of so many other denizens in the same middle ground: inscriptive

media all inhabit that turf, including plenty of forms that are even less familiarly "media" than the quartos and folios that concern traditional bibliography. A ten dollar bill is neither pure text nor untext, for instance, since its meaning (its value) as ten dollars arises amid consensual circulations of related tender: a bill is not "only" its value but is also not "only" a piece of paper.[3]

While bankers and computer engineers pondered the meaning and/as material of checks, computer punch cards broadly raised many of the same questions that draft cards did in the mid- and late 1960s.[4] Both kinds of cards contained a variety of information and both were also instrumental to the systems of which they were a part. If punch cards seem more literally instrumental—more on the bookmark end of things—than draft cards do, ironically punch cards also seem more pure, or self-identical with the information they contain. The contradiction arises because the systems that used them were big pieces of machinery and so much of the information they contained could only be "read" by those machines. The cards might have columns, numbers, words, or icons printed on them for human readers, but their primary instrumentality had to do with their mysterious (that is, illegible) patterns of holes.[5] While it is unlikely that anyone asked at the time how punch cards worked like texts, using them raised this question by proxy.[6] Users had to know what code the cards required, and how to punch them or have them punched to contain programs and data. They had to know up from down, front from back, used from unused, correct from incorrect—all questions that concerned the cards as meaningful instruments and the patterns of holes punched in them as a form of—or a form resembling—semiotic presence.

Both kinds of cards also came with injunctions against destruction, made explicit by an act of Congress in the case of draft cards, and made manifest by printed warnings in the case of many punch cards. "Do not fold, spindle, or mutilate": These exact words may not always have appeared but they enjoyed currency as a catchphrase, part of the social meanings of computer technology and its applications.[7] Several months before O'Brien burned his draft card in Boston, the student newspaper in Berkeley joked that "lesson one" for every new student at the University of California was "not to fold, spindle, or mutilate his IBM [student registration] card." Another system meant to pair bibliographic bodies and eighteen-year-old human ones, but that semester the Free Speech Movement effectively turned the university's punch cards into tokens of authority, to be rebelled against, altered, and subverted: one card puncher was used to make holes that spelled out "FSM" and "STRIKE."[8] Like O'Brien burning his draft card, these radicals subverted bib-

liographic norms and expectations as an exercise in free speech. They took cards that were intended to work in one way and made them work in another.

Many people—on the Left, at least—simply stopped drawing distinctions between one card-enabled system and another. Whether the cards registered draftees or pupils, they helped "the system." Lewis Mumford (1970, 183) called it "the Pentagon of Power," and reflected that "today the increasing number of mass protests, sit-downs, and riots—physical acts rather than words—may be interpreted as an attempt to break through the automatic insulation of the megamachine, with its tendency to cover its own errors, to refuse unwelcome messages, or to block transmission of information damaging to the system itself." Mumford's metaphoric megasystem suggests the broad cultural currency that "cybernetic anxiety" had come to enjoy by the late 1960s.[9] Anxiety, like alienation, was one response to dehumanizing bureaucracies with dehumanizing machines. Digital systems in particular seemed poised not merely to displace humans—as industrial automation kept doing—but finally to erase the very distinction between human and machine. This was the slippery slope toward "the posthuman" condition, as it has been called by N. Katherine Hayles.

The point of these examples from the 1960s is to suggest that bibliographic questions persist in unlikely places and can have unusually broad implications. Arguments over both draft and computer cards revealed unsettled and yet widely held assumptions about the ways that meaning is authored as well as conveyed on paper. These assumptions served to underpin the normal (that is, the systemic) uses of the cards, and consequently to circumscribe their adaptive reuse or destruction as subversive. Yet because they remained unsettled, the same assumptions helped form ripples in the social order, bubbling to the surface in contests to define the contours of the U.S. public weal, in arguments over free speech and attempts to parse the difference between speech and nonspeech, between Mumford's "physical acts" and "words."

One of the core propositions of the preceding chapters has been that unsettled assumptions like these become more unsettled, or at least more *evidently* unsettled, by new media. Put another way, new media can be potent, embodied versions of unsettlement. To be sure, there was nothing new about draft or, strictly speaking, punch cards in the 1960s. The most celebrated new medium of the decade was color television, the commodity's commodity. Color broadcasting systems had been developed in the 1940s and 1950s, but RCA publicly launched its system with great fanfare at the 1964–1965 New York world's fair.

At first glance, television in black and white or color has nothing whatsoever to do with bibliography since television is noninscriptive: a television broadcast has no body—unless or until it is taped. But as Philip Rosen (1994, 225–234) has shown in his analysis of NBC's coverage of John F. Kennedy's assassination, part of the defining portentousness of television arises from its use of inscriptions. Part of the anchorperson's presentation of that historic event as both self-evidently historic and uniquely available to viewers was a tension between the anchors' live narration and the still photographs they held up to the camera, the wire-service reports they read, the audiotapes they played, and particularly—more than two hours after the event—the first in-the-can footage from Dallas, taken earlier that November afternoon in 1963. The live broadcast was chock-full of inscriptions. Rosen calls this the "subtextual drama of the medium's struggle to depict itself," which was played out in the newscasters' and the network's evident discomfort at "the structuring absence of images from the key scenes of the motivating action" (228, 229). Even this preeminent noninscribed media event, the emergency broadcast, must be involved with and defined by its manipulation of as well as its partial yet repeated distinction from inscriptive forms.

I want to be clear that the media of the 1960s, new or old, in this respect have much in common with the new medium of recorded sound that I have described at length. I argued most explicitly in chapter 1 that when recorded sound was new, it was in some ways experienced as party to the existing, dynamic logics of writing, print media, and public speech, the nexus of so many open questions I have here called bibliographic ones, because I started with meanings authored and conveyed on paper. The new medium came to make sense only when its demonstration to and subsequent use by early audiences helped to construct a coincident yet partly contravening logic for recording—a logic that soon became self-evident, and thus came to seem intrinsic to phonographs and phonograph records. Furthermore, I suggested that the implications of the emergent logic for recorded sound extended far beyond the eventual formal conventions of the medium. Like the bibliographic questions that got so tangled up with the First Amendment in the 1960s, questions concerning the new and curious inscriptive qualities of sound recording were similarly entangled. Instead of free speech, recorded sound helped to broach questions about speech itself, the means and meaning of its selective preservation, which in turn helped to broach questions about the scope and character of U.S. public life and public memory.

This chapter concerns media that are more familiarly new than the new media of 1878–1910 or color television. I want to address digital networks as new media, both be-

cause they continue to dominate present thinking about media in general and newness in particular, and because they proffer an important instance of—very broadly—bibliographic questions and entanglements. Digital networks are many other things too, to be sure, and this chapter cannot presume to offer anything like a full accounting or complete history. My purpose instead is to strike a few notes of comparison between apples and oranges—between acoustic recordings and texts on the Internet—because I think those notes can lead, at least speculatively, to a better understanding of the ways that media accumulate the power that they do.

The differences between early phonograph records and the Internet are as obvious as they are undeniable: analog and digital, mechanical and electronic, consumer good and communications medium. Nevertheless, as new forms of inscription these two new media shared similarities. If sound recording helped to call the mutual meanings of print and public speech into question, electronic documents have raised questions of a related sort. Distributed digital networks and the texts they make possible have emerged amid an existing textual economy, a world and workplace powerfully self-constituted according to the logic of contemporary media: print publication, broadcasting, Hollywood, and the record labels, but also punch cards, printouts, and paperwork. Experiences with digital networks have helped to construct a coincident yet contravening logic for digital texts, partly in response to material features of the new medium, and partly in response to the hugely varied contexts of their ongoing reception and development. As digital media continue to change, the logic of electronic texts of course remains inchoate, although certain elements of that logic have already become conventional, and thus self-evident or transparent and difficult to see.

Like television, digital media at first blush have little to do with inscription. Inscriptions are both material and semiotic, and yet digital texts can seem strangely immaterial or disembodied. Like so much online, they are often thought of as "virtual" because they are so elusive as physical objects. No Web page would exist without a vast clutter of tangible stuff—the monitor on which it appears, but also the server computer, the client computer, the Internet "backbone," cables, routers, and switch hotels—but it is nonetheless strikingly intangible. What is it? Where is it? In this way, electronic texts seem to gesture toward the pure-text extreme dismissed as so unrealistic by Chief Justice Warren in relation to O'Brien's draft card: a digital text, according to this line of thinking, is only meaningful *as* the information it contains, so it is no problem—and no wonder—if the digital text seems to have no body.

Many scholars have recently pointed out the pitfalls of such reasoning, arguing for and exploring the varied materiality of digital texts. The best among them call for a "bibliographic/textual approach" to consider elements like "platform, interface, data standards, file formats, operating systems, versions and distributions of code, patches, ports, and so forth," because *"that's the stuff electronic texts are made of."*[10] But what sort of stuff is that? Even the most astute and exacting critics of cyberculture tend to signal a certain ambivalence about the bodies that electronic texts have, judging at least from the frequency with which the word *material* appears between scare quotes. Lev Manovich (2001, 45, 48) writes that the "basic, 'material' principles of new media [are] numeric coding and modular organization," and that hardware and software have "'material'" as well as "logical principles." Similarly, Mark Poster (2001, 77) chides that "the impact of technologies is never a linear result . . . of their internal, 'material,' capabilities."[11] The quotation marks around *material* serve obliquely to interrogate the claims being made. Both critics imply that there is no putting a finger exactly on the matter at hand, if logic is logic, but material is "material."

This is an ontological problem as much as a semantic one, a quandary over what a digital text fundamentally *is,* which has led, it turns out, to some productive wrangling over what a nondigital text fundamentally is by comparison. "What Is Text, Really?" Steven J. DeRose, David G. Durand, Elli Mylonas, and Allen H. Renear all wanted to know in the *Journal of Computing in Higher Education* (1990). "What is a text?" asked Espen J. Aarseth in *Cybertext: Perspectives on Ergodic Literature* (1997, 15) and Johanna Drucker in a review titled "Theory as Praxis: The Poetics of Electronic Textuality" (2002b, 685). "What is *text?*" Matthew Kirschenbaum insisted in *Unspun: Key Concepts for Understanding the World Wide Web* (2000, 127).[12] Critics always ask these questions rhetorically, in order to answer them, although their recurrence and persistence as questions in these contexts demonstrate an ongoing negotiation over the very subject of humanities computing. And in the much broader world of computing generally, similar negotiations abound. Any new editing, publishing, or scanning software, every new markup grammar, plug-in, or Web applet proposes a new idea of what electronic text actually is. The question is not asked as explicitly, but the answers are all there, and many of them are for sale. Any electronic text "embodies specific ideas of what is important in that text," and thus embodies specific ideas of what electronic text can be (Sperberg-McQueen 1991, 34). The high-stakes "drama of the medium's struggle to depict itself," as Rosen puts it, gets worked out both offscreen and on-screen, amid and as part of the political economies of knowledge work.[13]

Taking a page from DeRose et al. (1990), this chapter and the next explore digital text partly in relation to nondigital text. Just as the tinfoil records of 1878 and the wax cylinder records of 1889–1893 are knowable largely through the newsprint records that also helped to define them by contrast, electronic texts are knowable partly through and by contrast to "the social life of paper" and partly in relation to filmic, magnetic, and other inscriptions that are not paper based.[14] Just as chapter 1 looked at phonographs before they were even practical, everyday devices, this chapter considers digital text from the moment before digital networks were practical and familiar, as it appeared on the ARPANET, the precursor to the Internet (put together by the Advanced Research Projects Agency or ARPA) in 1968–1972. Structuring my inquiry this way allows me to sidestep the ontological problem of what electronic text *is* in order to try to understand how the *question* of what electronic text is ever occurred in the context of digital networks in the first place. What kinds of experiences might early users of digital networks have had that implicated those networks as relevant to issues of bibliography at all? What were the contexts of use and usability on the emerging network? Only by tackling these issues first will it be possible to tease out some of the sprawling implications that the relevant bibliographic questions may have. The following chapter takes a stab in the latter direction by focusing on the World Wide Web today.

Sources for my immediate project are admittedly scant. Electronic records from the ARPANET era have rarely survived or have been "ported forward" to more recent technology in such a way that their ARPA-ness, their telling archival character or provenance, whatever it might have been, has become obscure.[15] Published sources from the ARPANET era are also scarce, particularly when compared to the bubble of attention that Edison's first phonographs received. In the popular press, for example, the *New York Times* only mentioned the Internet once before 1988; then too, there was little relevant trade literature to speak of, and scholarly publication in the nascent field of computer science was only beginning to gain momentum.[16] As an agency within the U.S. Department of Defense, ARPA and its Information Processing Techniques Office (IPTO) didn't publicize its activities. Even the IPTO's contract bidding process was—an oxymoron—selectively public.[17] The dearth of electronic sources in particular presents a problem, although it also addresses one of the central points of my argument about new media and media history. Much of the potential standing that electronic texts have as evidence is an outcome, not a precondition, of the logic that digital media are coming to possess specifically in relation to writing, print, and other media.

Keyword: Document

When J. C. R. Licklider went to the Pentagon in fall 1962 to head APRA's new IPTO, he took a leave of absence from Bolt, Beranek, and Newman (BBN), a Cambridge, Massachusetts, firm of engineering consultants who specialized in acoustics. Over the next few years, Licklider would help to envision and initiate the ARPANET project, and then BBN would be hired under his successors at the Pentagon to build the network's core Interface Message Processors (IMP). In the meantime, one of the things Licklider left hanging at BBN was a consulting gig for the Council on Library Resources, Inc., funded by the Ford Foundation. Though removed from the day-to-day research, now directed by a colleague at BBN, Licklider kept track of the project and eventually authored its report. As the Council on Library Resources noted, "The 'research on concepts and problems of libraries of the future' continued under [Licklider's] general direction in his absence," while he was "on a special assignment for the Department of Defense." The council got its final report in January 1964, and The MIT Press published *Libraries of the Future* in 1965.[18]

The impetus for the libraries project arose from what Carolyn Marvin (1987, 59) has called a "digital" view of information, although the term *digital* was not used as freely at the time. The digital view assumes that "information is a definite quantity increasing daily," as opposed to what Marvin identifies as an "analog" view, characterized by the assumption that information is "a constantly shifting and repatterned feature of every environment, past or present, a transaction between knowers and what is known."[19] The digital view of information has Enlightenment roots, but it emerged with force in the post–World War II, post–Manhattan Project era, as part of a widely shared anxiety about the continued efficacy of science in U.S. life—an anxiety that only intensified with the Soviet launch of Sputnik in 1957. In particular, pundits worried, science faced an impending bibliometric crisis: too many publications. Increasing specialization and ever more research meant that scientists—and all of civilization by extension—were being buried alive by the accumulated mass of paper. If left unchecked, Mumford (1970, 182) summarized, the sheer "overproduction of books will bring about a state of intellectual enervation and depletion hardly to be distinguished from massive ignorance."[20] University and other research libraries were already "becoming choked from the proliferation," according to the Council on Library Resources (Licklider, v). So the council hired Licklider and BBN to work on the problem.

The most famous statement of the impending bibliometric crisis is Vannevar Bush's "As We May Think," which had appeared in the *Atlantic Monthly* in 1945.[21] The article offers a

meditation on the future of compiling and consulting the huge mass of accumulated human knowledge, which Bush variously calls "the great record," "the total record," and "the common record" of humankind. Bush's formulation is famous because he describes a futuristic solution to the problem: a method of storing and sorting information that is modeled on human consciousness rather than bureaucratic filing systems. It is an imagined hypertext, before the term *hypertext* was coined and any of the relevant digital technology existed. He suggests that documents might best be organized not by "artificial" indexing systems with their rigid "paths" and cumbersome rules but by a more natural form of "associative indexing," working in the manner of the "intricate web of trails" that connects related thoughts in the brain. To this end, Bush posits a device, "a sort of mechanized private file and library" to serve as an "intimate supplement" to its user's memory. Dubbing it the *memex,* Bush envisions "a desk [that can be] operated at a distance," with screens on top and microfilm contents inside to be somehow selected, consulted, annotated, and then joined or tied at will into multiple associative "trails" for future reference.

Bush's presumptive "we"—"As *We* May Think"—like his preoccupation with "the common record," signals a media public of the sort I explored in chapter 1. Like the exhibitors of tinfoil phonographs, that is, Bush's memex posits the shared ownership of a public record, presumptively "owned" by "us" in the commonsensicality of its own production, preservation, and potential retrieval for a shared public good. Licklider tackled the libraries of the future project with all of these same assumptions in place. What he added was his own futurism, an exacting interest in "Man-Computer Symbiosis," and an extraordinarily astute analysis of libraries and their role in the production of knowledge.[22] Though he claimed not to have read Bush's "pioneer article" until completing his own work, Licklider acknowledged its indirect influence, "through the community," and dedicated *Libraries of the Future* to Bush (xiii). Licklider's volume is much less well-known than Bush's article, but it offers an opportunity to gauge the ways that computers and texts could be thought together in the moment just before digital networks became an accomplished fact as well as a new medium.

The future Licklider takes as his point of orientation is the year 2000, and the libraries he proposes are what he calls "procognitive systems." In order to envision these future procognitive systems, Licklider must reject many of the "schemata" by which traditional libraries are understood as well as many of the schemata by which computers were understood in the mid-1960s. He is in favor of the page, for instance, as a schema for the display of information ("As a medium . . . the printed page is superb" [4]), but dismayed by its "passiveness" as a means of long-term information storage. Books have even less to

recommend them, and library stacks of books still less. Likewise, Licklider is in favor of some of the features of computing, like random-access memory (RAM), but not of the existing technological limitations or the mind-set that associates computing with writing a program and delivering a stack of punch "cards to a computer center in the morning, and picking up a pile of 'printouts' in the afternoon" (8). By carefully rethinking libraries and computers in such schematic terms, Licklider arrives at a wishful future in which researchers sit at consoles or terminals, typing on keyboards and looking at screens, connecting to and interacting with digital systems to query, search, and retrieve information.

Licklider's procognitive systems are amazingly prescient. Computing in "real time," when a user and computer interact through keyboards and screens, was all but impracticable in 1965, known primarily to those few who worked on advanced systems for the military. Drawing on his experiences with computing at MIT, where real-time and time-sharing (multiterminal) computing were both under continuous development, Licklider had settled on the same man-at-his-desk schema that had shaped Bush's memex and that would so influence the graphic user interface—with its metaphoric desktop—of personal computing more than a decade later.[23] Licklider envisions a year 2000 in which "the average man" owns or rents his own terminal and uses it to connect to a network. "In business, government, and education," he speculates, "the concept of 'desk' may have changed from passive to active: a desk may be primarily a display-and-control station in a telecommunication-telecomputation system—and its most vital part may be the cable ('umbilical cord') that connects it, via a wall socket, into the procognitive utility net" (33).

The image of man and his umbilical cord is a striking one, but what kind of actual or bibliographic body will "the body of recorded knowledge" (6) have when Licklider's procognitive systems become the libraries of the future? Licklider and his research team recognize some limits: "We delimited the scope of the study, almost at the outset, to functions, classes of information, and domains of knowledge in which the items of basic interest are not the print or paper, and not the words and sentences themselves—but the facts, concepts, principles, and ideas that lie behind the visible and tangible aspects of documents." By focusing on information that exists "behind" the visible and the tangible, Licklider and his team agree to set aside all works of art, graphic and literary. Instead, they focus on cases where they suppose that information can be separated from its natal body "without significant loss": both art and literary criticism, perhaps, as well as "most of history, medicine, and law, and almost all of science, technology, and the records of business and government" (2). The information "corpus" so disassociated from pages, books, and shelves would be introduced to a new body, an advanced, "processible" mem-

ory system or binary storage medium, as yet undeveloped. That memory, Licklider estimates, would have to have at least the capacity for 5×10^{15} bits, where each alphanumeric character in the corpus is represented (conservatively) by 5 bits, taking 5 cells of binary storage space (15–20, 63–64).

Without a specified core memory, Licklider's procognitive systems lean toward the pure text ideal hinted at in the O'Brien decision, with one notable addition. Licklider compares his procognitive systems to a world in which a "document [can] read its own print" (5–6). The systems would be able to answer questions as well as retrieve documents, because they would "be able to 'read' and 'comprehend' the documents themselves" as well as the tags or catalog information associated with them (153). So Licklider's electronic documents are self-identical and self-reading. They *are* the information they contain; yet they also consist recursively of representations, "smart" versions of themselves and the paper documents they reduce to facts, concepts, principles, and ideas.

By *document* Licklider explains that he means a document "type," not what he distinguishes as individual document "tokens." A document token is a unique and physical copy of a document, whereas a document type is a whole edition of like tokens from which any representative may be selected to have its alphanumeric characters spelled out into the system (13). Any variation among document tokens is, like typography, assumed to be trivial, dispensable, or what bibliographers call "accidental" or nonsubstantive. But if textual variants among tokens are unimportant, Licklider still wants his procognitive systems to retain a number of concepts or schemata relating to variations among document types, like genre: monographs are different from articles are different from reviews. And he wants to retain schemata relating to the hierarchical parts that document types have, such as sentence, paragraph, chapter, and volume as well as author, title, abstract, body, and footnote (7).

Libraries of the Future contains a brief survey of the relevant research on information storage, organization, and retrieval along with a section on the linguistic analysis that might be required in the "eventual machine 'understanding,' of natural-language text" (131). Licklider himself envisions a sequence in which the "natural," technical language of science would undergo "a machine-aided editorial translation" into an unambiguous or "ruly" English, which would in turn undergo "a purely machine transformation" (that is, scanning) into the "language(s) of the computer or of the data base itself" (89). For the present, that means that the proposed transmigration of information from natal to self-reading body hinges on the conversion of documents into so many alphanumeric character strings (as well as some sort of unspecified "adjustment" for pictures and other nonalphanumeric

contents) that are retrievable and readable by as well as via the procognitive systems. Alphanumeric characters are essential, but actual letterforms don't matter, even though ironically the document at hand, *Libraries of the Future,* depends on font changes to signal shifts between Licklider's analysis and the examples he gives of hypothetical interactions with both a local computing system and the procognitive one. Each system is represented by a different font, and both are different from the prevailing typeface of the book. "Are you J. C. R. Licklider?" the local computing utility asks Licklider in a typewriter font. "I type 'y' for yes," Licklider narrates entirely in the dominant font of the book. When he types "Procog" to log on, the system responds, "You are now in the Procognitive System" in a small san serif typeface (47–48).

Licklider gives an illustration of a procognitive system in use that is strangely ambivalent in its futurism. He uses a procognitive system from the future, but to pursue a question he has in the present, explicitly as if he had "available in 1964 the procognitive system of 1994." He asks the system about "the prospect that digital computers can be programmed in such a way as to 'understand' passages of natural language" (46). The procognitive system searches ten thousand documents for relevant material, then presents Licklider with citations and abstracts, and has full-text versions "available in secondary memory," if he wants them. His online session is dated 14:23 November 13, 1964, and the bibliographic citations he retrieves refer to articles from 1961 and 1963 that appear in his own bibliography. In this respect, even though the paper-based *Libraries of the Future* is not a digital or self-reading document, it is represented by Licklider's example as a self-writing one. The conceit of self-writing whereby *Libraries of the Future* imagines the future research that leads to *Libraries of the Future* offers an echo of the recursive logic that any system of self-reading documents must possess as well as the reflexivity or self-study that would be offered two years later as one of numerous reasons for building the ARPANET.[24]

Many aspects of Licklider's future do resonate with actual developments in computing over the decades that ensued. If the desk schema, query, search, and network interactions all make sense to today's computer users, many of Licklider's more specific ideas about digital texts also make sense in hindsight according to later developments. Licklider's wishful 5-bit alphanumeric code can be recognized in the 8-bit ASCII standard adopted by the ARPANET in 1969. His wishful commitment to genre can be recognized in today's document-type definitions or markup grammars, which text encoders construct for different classes of documents. And his wishful commitment to the hierarchical parts that documents have can be recognized in the principle of defining texts as ordered hierarchies of content objects on which markup strategies are premised.[25] Tak-

ing *Libraries of the Future* on its own circa 1965 terms, however, helps to index more about digital inscriptions than the forms they have recently taken or the theorizations they may lately have provoked.

In particular, Licklider's volume demonstrates both the pertinence and complexity of the term *document*. Like the related term *record* in 1878, the term *document* in the 1960s was nothing new. As a noun, it hails in its present form from the eighteenth century, when, according to the Oxford English Dictionary, it referred to "something written, inscribed, etc., which furnishes evidence or information upon any subject, as a manuscript, title-deed, tomb-stone, coin, picture, etc."[26] As Licklider uses the term, documents have something of a synecdochic function: documents are not unique artifacts (his "tokens"); they are representative ones. In addition to furnishing evidence or information on their subjects or contents, they furnish as well as delimit evidence about the whole edition they represent. (They importantly delimit evidence because they are based on assumptions about which details matter and which do not.) Moreover, since they are called "documents" when they are shelved on a bookcase and also once they are entered into the procognitive system, their status is doubly synecdochic online: electronic documents are representative-substantive copies, simulacra, of already representative-substantive forms. They furnish as well as delimit evidence about types that furnish as well as delimit evidence about tokens.

Documents in *Libraries of the Future* may be digitized, but they cannot be de novo digital creations. Licklider proceeds according to an unstated distinction between documents and what Matthew Kirschenbaum (2002) has called "first-generation electronic objects." Licklider's electronic documents are third-generation, edited from natural language into "ruly" language and then processed into machine code. In Licklider's volume, first-generation electronic objects appear only as the prompts, headers, queries, results, and other online interactions that are represented on the procognitive screen or—one of the hardiest schemata of computing in 1965—as printouts on sheaves of paper. Put another way, there is always an inside and an outside to documents within the Licklider system. Documents are distinguished from programs as well as "raw data, digested data, [and] data about the location of data" (Licklider 1990, 29).

Today, the meaning of *document* is much broader than this and at least as complex, to judge merely from the "My Documents" folder on any personal computer running Windows, the ".doc" files created and accumulated by Microsoft Word, or the sophisticated software packages, hardware alternatives, and consulting services available for document management.[27] Documents today may be created digitally, not just digitized. Licklider's

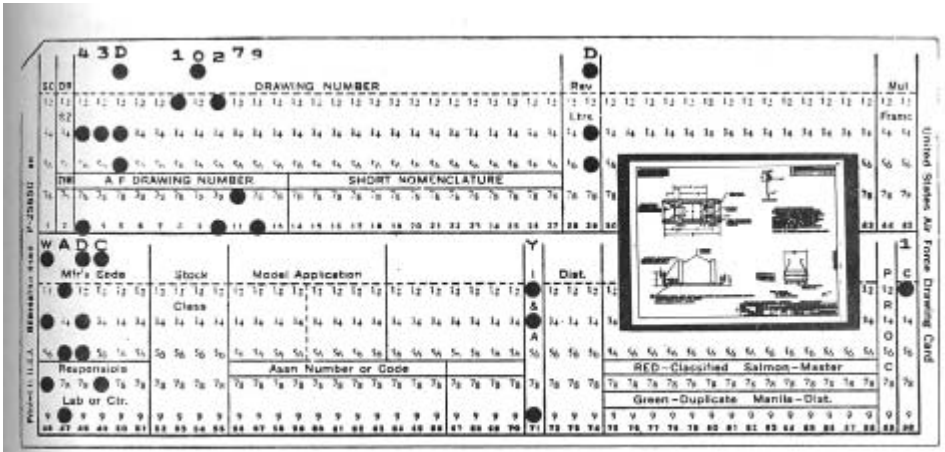


Figure 1. Remington Rand Format

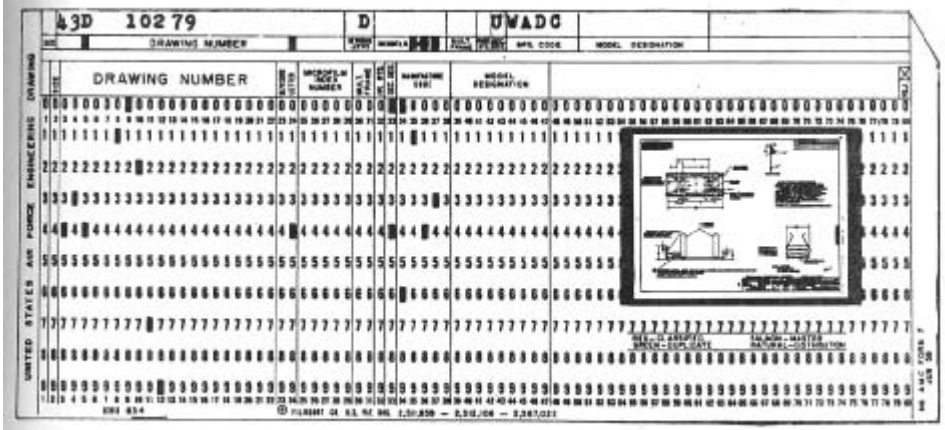


Figure 1A. IBM Format

Figures 3.2a and 3.2b Microfilm aperture cards and reader, 1960. (*Source:* Charles Babbage Institute, University of Minnesota.)

usage still lingers, however. It helps to explain why the Xerox Corporation trademarked the slogans "We Document the World" (1988) and "The Document Company" (1991), with the latter mark still in use. Indeed, Licklider's electronic documents in some respects resemble "smart" or self-reading photocopies. Like a photocopy, his digitized document furnishes as well as delimits evidence of the preexisting paper document it represents. Likewise, just as there is no de novo digitized document in Licklider's scheme, there can be no de novo Xerox, a photocopy made wholly without reference to a preexisting document. The Xerox Corporation's slogans, of course, have much to do with its long-lived ambitions to move beyond photocopiers and toward the synergies of the modern office environment. As early as 1966, the company needed engineers and other specialists to help develop "document management and computer-related systems"; help-wanted advertisements started to include "document management" as a discrete object for R & D around that time.[28] Among the most inventive systems for document management (not

by Xerox) were the microfilm aperture card systems, a primitive database of images controlled on computer punch cards and read on a memex-looking desk.[29]

The present meanings of *document,* to stay briefly at this suggestive, semantic level, are no more singularly the result of digital media than the enlarged meaning of the word *record* in 1878 was the singular result of Edison's invention. Phonograph records added one specific meaning to a word already in the process of redefinition; they emerged from and engaged a conceptual field already subject to stress. Newly applied to tinfoil sheets as well as differently and newly applied to people (the politician's record, the ballplayer's record, and so on) the word *record* as it was used around 1878 hinted at what seemed important to record and important about recording, while both forms of importance appeared keenly in question amid the varied cultural politics of the post-Reconstruction era. The term carried with it and was reflexively defined by that portentous sense of abstract public or "common record" to which Bush later appealed in his imagination of the memex, although the public of 1878 was hardly more certain in its extent than that of 1945 or of today. Likewise, digital media have added specific—though notably loose—meanings to the term *document* as part of an already ongoing redefinition that remains complicated by and within a sense of abstract public and common record, of what seems important to document and important about documents.

"Importance" in this context alludes to that nexus of cultural values attending and reflexively defined by overlapping and dynamic media publics. Those values may be glimpsed in other uses of the term *document,* as in such varied productions as "documentary" editions and films. Documentary editions are the work of historians and textual critics who locate, authenticate, select, annotate, transcribe, and publish documents for use by students and scholars in the future; for example, *The Papers of Thomas A. Edison, The Margaret Sanger Papers,* or *The Mark Twain Papers.* They are documentary in the sense that they assume the merits and efficacy of historical and textual editing as a professional practice, of the work that humanists do by reading, assessing, and citing primary source material to support narratives about the past. Documentary editions supplement the "database logic" of archives, Lev Manovich (2001, 218–232) might say, promoting documents for their explanatory power (Manovich's "narrative logic") according to the norms of the historical and literary-critical professions in addition to the broader cultural economy that supports those norms in their varied attentions to Edison, Sanger, Twain, and the like. Documentary films by contrast are documentary in the sense that they assume the merits and efficacy of the filmic medium to represent reality or present nonfiction. Both uses of *documentary,* like so many uses of *document* as a noun and verb, differently denote a con-

nection between texts and facts, where the facticity of texts inheres in their varied self-evidence as meaningful bodies that are also potentially authentic, original, unique, complete, uncorrupted, lasting, immutable, citable, or otherwise "true."

The related genealogies of *record* and *document* suggested here together point toward a history of the facticity of the modern text. Where has the present, if varied, self-evidence of texts come from? How have the meanings of textual bodies evolved as such, and according to which culturally and media-specific conditions? What would the history of bibliography look like in the broadest possible terms? Such a project would be in keeping with the "history of the modern fact" that Mary Poovey (1998) has discerned first in the construction of numerical certitude by Renaissance bookkeepers, as well as with "the invention of communication" that Armand Mattelart (1996) has discerned first in the work of seventeenth-century topographers.[30] It would have to include—even more self-reflexively—the genealogy of *text* in the postmodern or "antidisciplinary" sense that John Mowitt (1994) has dated to the poststructuralist movements of the same late 1960s that have concerned me here. And it have to would include an ample share of what is now called "material culture studies," a genealogy of cultural things and thingness.

One might begin this history of the modern text with the conceptual emergence of editions, and thence of modern bibliography, whether from the *pecia* system among Renaissance scribes (a division of labor among copyists) or the print shops of Gutenberg and his contemporaries. Or one might begin this history with the conceptual emergence of print as a fixed, authorizing, and authored medium, as with the British Stationers' Company and the Royal Society.[31] There are many other alternatives too. Wherever and whenever one begins, the printed book, like that later construction, "literature," is only one of its themes. Such a history must include the history of the footnote and photography, the imagination of the memex, the burning of draft cards, and the birth of digitization.[32] The varied means and subjects of inscription all offer themes, as do the varied political and cultural economies of making inscriptions public and consuming them. This whole history of the modern text cannot be my purpose here, yet it does offer a reminder that the histories of sound recording and electronic documents go together as one, not only as two.

1968–1972: Networking

The early architects of digital networks had to reckon with bibliographic questions on a number of different levels. The first, most obvious level was an immediate condition of bureaucracy. The ARPANET began with four nodes located at four universities that were

already research contractors for ARPA. Representatives from each of the first four sites met during 1968 to begin planning the network, and the minutes of these early meetings were copied and circulated to ARPA. Primary investigators brought and later sent their graduate students to these Network Working Group (NWG) meetings, and representatives from BBN, the contractor selected in January 1969 to provide the interface computers at each node, joined them in February. It wasn't until March 1969, apparently, that it occurred to members of the resulting and somewhat ad hoc group that they themselves should really begin to write things down. This is not to say that paperwork had not already been accumulating at precipitous rates, particularly in Washington, DC, just that the group itself, this nascent online community comprised largely of graduate students, had taken few pains nor realized the need to generate a collective memory of its own over which its members—or the Department of Defense, or anyone—might have bibliographic control.

At the same time that it picked BBN to start construction of the IMP, ARPA had also awarded the Stanford Research Institute (SRI) node a contract to become the "Network Information Center" (NIC), a documentation site for the projected network. The resulting and initial efforts to document the in-progress ARPANET, by both the NWG members and the nascent NIC, together demonstrate an emergent interest in the genres and textual or bibliographic bodies appropriate to the new medium, an interest that would form the basis for an emerging "recursive public," in Christopher Kelty's terms (2005, 186), "a particular form of social imaginary through which [a] group imagines in common the means of [its] own association."

In April 1969, the NWG announced a new genre for the new medium: the request for comment (RFC). The name was an appropriation of the Pentagon's habitual requests for proposals, or RFPs, which solicited proposals as part of its contract bidding process. Put together, RFCs were "a set of notes" in which the NWG discussed the requirements for the network and established its initial protocols. There had to be a protocol for establishing protocols, however, so graduate student Steve Crocker tried to define the RFC genre in RFC 3 (April 9, 1969): "The basic ground rules were that anyone could say anything and that nothing was official." In that memo, he specifically encourages "philosophical positions without examples or other specifics, specific suggestions or implementation techniques without introductory background explication, and explicit questions." As Crocker explains, "These standards (or lack of them) are stated explicitly for two reasons. First, there is a tendency to view a written statement as ipso facto authoritative, and we hope to

promote the exchange and discussion of considerably less than authoritative ideas. Second, there is a natural hesitancy to publish something unpolished, and we hope to ease this inhibition." RFCs are writing that seeks to be less written, publications that seek to be less published, and records that might be less than permanent. Crocker apparently believed that RFCs were "temporary and the entire series would die off in a year or so once the network was running." He thought that the online community (his "we") would eventually have the networking protocols established by the RFCs, but it would have no collective memory of its own apart from their implementation.[33] For the sake of clarity, Crocker proposed, "Every NWG note should bear the following information" in its heading:

"Network Working Group"
"Request for Comments:" x,
where x is a serial number
Serial numbers are assigned by Bill Duvall at SRI
Author and affiliation
Date
Title. The title need not be unique.

If Crocker was imagining a new form of document, he was still constrained by the existing technology. Originally RFCs were memos distributed on paper. At the outset, six copies went by mail, one each to Washington, the four nodes, and BBN, where further copies could go by hand. Within months, RFC 10 (July 29, 1969) updated the distribution from six to nine paper copies and gave mailing addresses. Subsequent RFCs enlarged the mailing list still further, and refined the processes for numbering and issuing them. RFCs circulated the parameters as well as the contested protocols of circulability, their own circulability as texts as well as the future circulability of packets, the unit of data transmission across so-called packet-switched networks like the ARPANET and its successors. As Kelty puts it, standards and protocols were "bootstrapped, first through a process of writing RFCs, followed by a process of creating implementations that adhere loosely to the rules in the RFC, then observing the progress of implementations (where they differ from the RFC, how many people are using this implementation, and other less obvious criteria), and then rewriting the RFCs so that the process begins all over again" (2005, 198). When early RFCs were composed and stored online, like RFC 2, they still had to be printed and distributed on paper, because the digital network as such did not yet exist.[34] Even after the first nodes were connected, when the NIC made its first

important plug for online, networked documentation, it did so in terms that stressed the importance of paper: "Particularly useful would be on-line documentation. It is suggested that each site have an identical hard copy device (e.g., a line printer) suitable for reproducing documents."[35] Online documentation was only going to work if every site could download to paper. Like draft cards, documentation was for keeps, and in 1970 keeping still suggested "hard" paper copies—the NIC's printed "reproductions"—stored in local files.

The RFCs work as a form of documentation for the network and thus as a performative self-iteration of themselves as networked information. (The series continues today.) Some RFCs stand as protocols, rules for building and using the network—they articulate Telnet, FTP, and so on; others simply document the refinement, existence, and accumulation of standards and protocols. In February 1971, RFC 100 was issued as "a guide" to the series as it existed then. Its author, Peggy Karp, categorized the RFCs by subject, indicating "whether the notes are current, obsolete, or superseded," and "for historical reasons," gives "a brief summary" of each RFC "relevant to a particular category." Further efforts to control the accumulating information followed.[36] Eventually, RFCs would be published and stored as public access files in ASCII, and an electronic message was then sent around letting subscribers on the network know that they were ready for download. After more than a decade, RFC 825 (November 1982) would explicitly formalize the "rules" for RFCs that had emerged as a result of "the constraints of a wide variety of printing and display equipment" across the network. If RFCs were going to be accessible to all, they had to be in ASCII; they had to consist of fifty-eight-line pages made of seventy-two-character lines, and "No overstriking (or underlining) is allowed."

Despite formal properties at least as rigid as those that define a sonnet or a villanelle, the RFC has been celebrated as an inherently democratic form. Like e-mail in 1972 and Usenet groups in 1979, that is, the RFCs are part of an already often-told history of the Internet that emphasizes the network's grass roots, its community-based self-definition of the communication protocols that include multiple layers of technical specifications, but that extend as well to the conventions and popular accessibility of online discourse.[37] While in one sense there is clearly something tendentious about describing any element of an insular, "closed world," Pentagon-sponsored, R & D project as "democratic," in another sense there is also something a little banal about it in this instance, since all genres, as genres, are socially derived.[38] What remains critical to note is the way that ARPANET standards and protocols emerged in interlocking layers. The layer at which the RFC became defined as an ASCII file of a certain kind was in part a function of the material, tech-

nical conditions of the network—the wide range of output devices for printing and display that were getting hooked up. It was as if the number and metrical length of lines in the sonnet had been suggested by the diverse practices of Renaissance papermaking.

And if the heterogeneity of output devices for printing and display had certain ramifications for online, networked documents, so did the heterogeneity of input devices. In November 1970, when the ARPANET had grown to more than ten nodes, the NWG met in Houston at the Joint Computer Conference to discuss its use and further development. Participants worried that the ARPANET did not yet have "significant operational usage," and were agreed that the availability of documentation would offer a crucial encouragement to potential users. According to RFC 77 (November 20, 1970), they decided to concentrate for the moment on keyboard input devices with printouts rather than try to tackle both keyboards and graphic devices like light pens and mice with screen display output. But they also worried about the prestige of keyboards:

[Doug] Engelbart: Perhaps we should put off graphics several months so as not to delay typewriters. Typewriters are important.

[Edwin Meyer]: But would that be sufficiently impressive for DOD [Department of Defense] people?

Even limiting the discussion to keyboards meant that there were some thorny issues yet to resolve. Engelbart of SRI reported that the NIC was setting up a "dialogue system" that included an online documentation service, and might even put "a communication agent and technical liaison officer" at every node. One thing he needed to know in order to get his documentation service off the ground was what sorts of "console interaction" each site had.[39]

Keyboard-based console interaction came in four basic types. It could be either full or half duplex, and either character or line oriented. A full-duplex system can transmit and receive at the same time (like a telephone), while a half-duplex system allows transmission in one direction at a time (like a walkie-talkie). Character-oriented systems respond to characters as they are typed, while line-oriented systems respond to complete lines of type that end with an end-of-line command, like a carriage return. The system response in question is called echoing, or as Jon Postel understood it from Engelbart's discussion, "the situation where the console, a peripheral processor or some very low level software, echo[e]s the character [or line that is] entered." Different nodes on the network handled input in different ways; some relied on echoing at a local level, while others echoed remote input themselves. The difference had less to do with what one communicated to the

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please e-mail <special_sales@mitpress.mit.edu> or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Perpetua by Graphic Composition, Inc. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Gitelman, Lisa.
Always already new : media, history and the data of culture / Lisa Gitelman.
p. cm.
Includes bibliographical references and index.
ISBN-10 0-262-07271-8 — ISBN-13 978-0-262-07271-7 (hc : alk. paper)
1. Mass media—History. 2. Communication and technology—United States—History.
I. Title.

P90.G4776 2006
302.2309—dc22

2005058066

10 9 8 7 6 5 4 3 2 1