

Private Circulation

Print this document on 32lb paper

UNIVERSAL ACCESS TO HUMAN KNOWLEDGE

Digitized information, especially on the Internet, has such rapid turnover these days that total loss is the norm. Civilization is developing severe amnesia as a result; indeed it may have become too amnesiac already to notice the problem properly. The Internet Archive is the beginning of a cure—the beginning of complete, detailed, accessible, searchable memory for society, and not just scholars this time, but everyone.

— Stewart Brand

It would be a shame if brilliant technology were to end up threatening the kind of intellect that produced it.

— Edward Tenner

In the fall of 2007, while searching for more lucrative employment, I took a job for two weeks scanning books for the Internet Archive, a San Francisco-based nonprofit

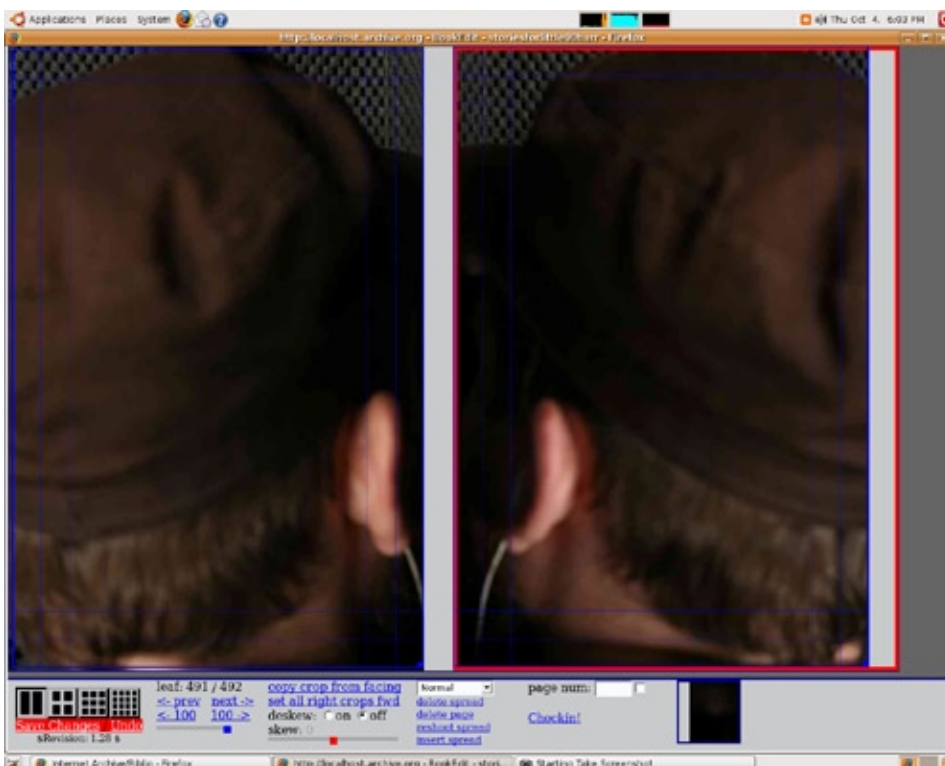
that is attempting to digitize every book in the public domain. Founded in 1996 by Brewster Kahle as a project to archive the entirety of the Internet every six months, the Archive's guiding motto is Universal Access to Human Knowledge. According to *Wired* magazine, it is "the world's largest online collection of free books, with nearly 350,000 titles and growing." During the two weeks when this essay was being written, the number of texts available through the Archive's homepage grew from 519,873 to 527,365.

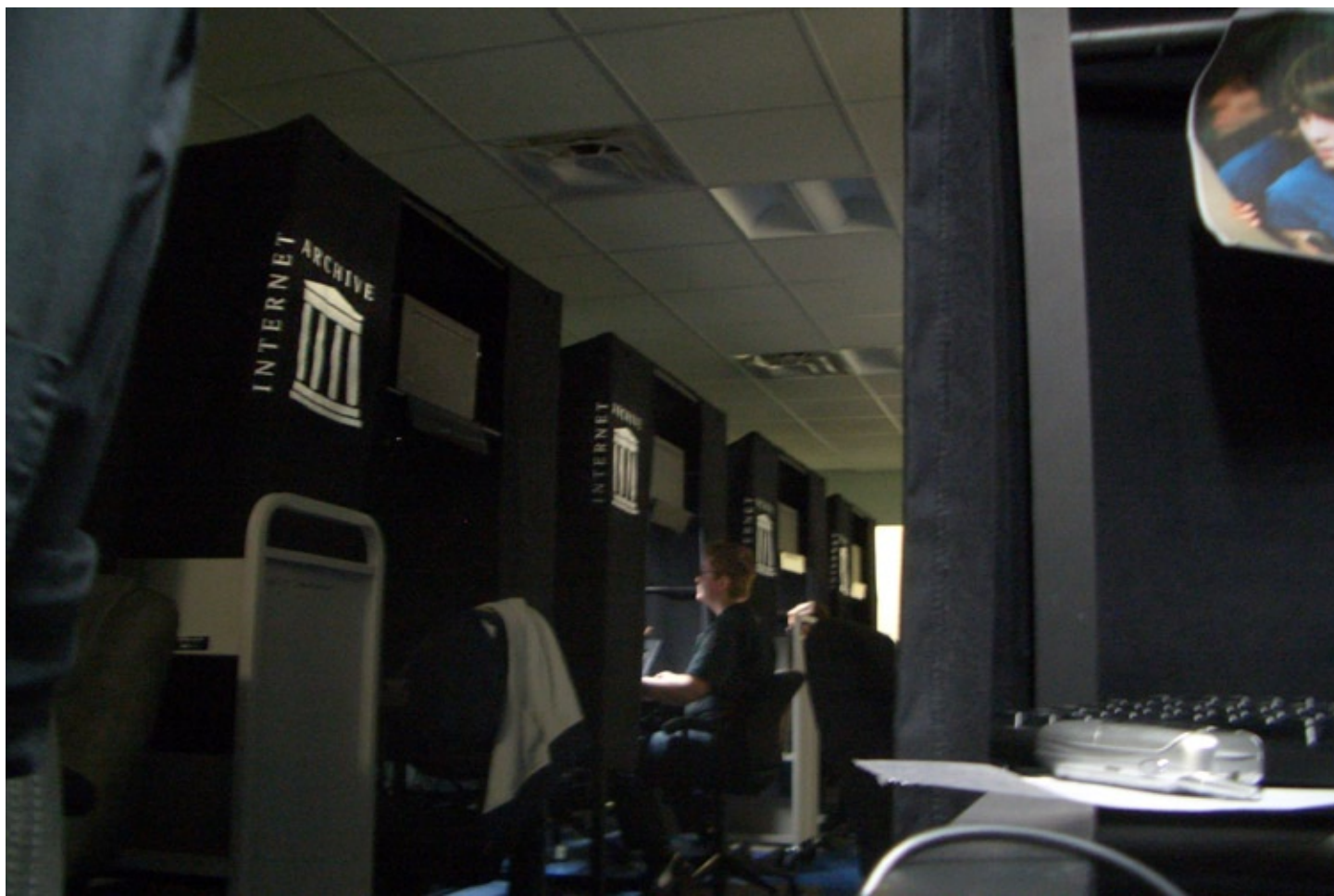
The Internet Archive's New York book-scanning satellite was located on West 53rd Street, directly across the street from the Museum of Modern Art on the top floor of the Donnell Library Center that closed in June of 2008. The Donnell operated for 52

years as one of the most popular branches in the city, and until its closure was home to a historic children's collection that included the original Winnie the Pooh dolls. The building will be razed to make way for an 11-story luxury hotel, though the library is slated to reopen in a section of the hotel building's first floor and a nebulous "underground area." (Library patrons and hotel guests will supposedly use separate entrances.) Come 2011, we'll see how much free public space will be allotted.

I arrived for my interview at 3 p.m., just as the night crew was settling in, taking over from the day crew. The crews were friendly with each other and chattered as they changed shifts. Abigail, my future supervisor, said that she liked my cover letter. In it, I had written, "I have a high threshold for repetitive tasks"—a line that I would soon come to regret. After several punctilious questions, she led me to what she called a "Scribe Station," a booth of apparatus cloaked by heavy black fabric bearing the Internet Archive logo (a Grecian-like temple). The Internet Archive's brand identity is peppered with historical references such as the Library of Alexandria, the Greeks, and the ideal of democratic access to information. A more apt metaphor for the drudgery necessary to accomplish such a project might be the toiling builders of other great public works such as the Great Wall or the pyramids.

The screenshot above shows the Quality Assurance screen of the Scribe software just before I rejected the page. I had protracted fantasies of importing fake books into the system, or inserting artwork into the blank pages of books, but never did.





There were two rows of these black-clad stations, nine in total, but Abigail briskly informed me that “only three or four were used during the night shift.” When the machines were in repose, their screens ran the classic Linux screensavers of bouncing cows and armadas of flying toasters, or were simply black. Occasionally, the mouse of an otherwise empty scribe station activated programs and ran diagnostics as if there were a ghost in the machine, although I later learned that they were being operated remotely from the Internet Archive headquarters in San Francisco.

The supervisor selected a book from the shelf of titles waiting to be scanned and handed it to me. She switched on a pair of tungsten lamps, carefully adjusted the two cameras (one for each page of the spread), and sat down at the station. At first I asked a lot of questions, but it soon became apparent that I would quickly become intimate with the intricacies of my task without recourse to communication. I watched as Abigail lifted the V-shaped glass platen using a foot pedal. Taking the book from my hands, she cracked it open, flipped it onto its front, positioned it on the base, and lifted her foot off the pedal. The glass platen wedged down into the open pages, conforming them to the angled glass. When she clicked the mouse, the apertures of both cameras snapped simultaneously. She once again lifted the platen with the foot pedal, flipped the page, and pressed the mouse; ditto. “You can also hit the spacebar,” she said, a seasoned expert. Then she demonstrated the second and final stage, QA, or, quality assurance, a last chance to check the margins of each page before the book was “republished” for posterity.

I took the job. I rationalized that I was a foot soldier with an important duty and that the repetitive motions would clear my head for more important work; not that I needed the rent money and had failed to quickly find more gainful employment. Besides, I like books,

Inside the New York satellite of the Internet Archive an employee works at one of nine Scribe stations during the night shift. The light from the background is a back-lit frosted-glass wall. I never found out what was on the other side, but imagine that it was a server room.



Top: The day manager's desk was empty at night and provided a surface for beverages, which were not officially permitted. From the red-lidded container we were allowed to take three candies per shift. This was the only job I have had where a coffee-maker was prohibited. Against the wall is a plastic-wrapped shelf of books that were next in line to be scanned. **Bottom:** Click on the image to watch the video, *Self-portrait Scanning Books*.

I told myself. As I soon found, however, many of the books that I proceeded to scan were of dubious quality, which raises the obvious question: If the material is junk, why and for whom are we creating this “universal archive”? Will such an inclusive depository ever be useful? In his article about a project to photograph every single house in America, George Pendle asks, “When does data become as inert as the reality it represents?”

There were no windows to the outside on the top floor of the building, which the Archive shared with an I.T. company. But on my way to the break room, walking down an institutional-blue hallway, I could see—through two vertical slits of glass—rows of tall, streamlined servers raised above the floor in their eternally lit, temperature-controlled room. When I think of the word *information*, I think of this room.

Three hours into my eleventh shift, I walked off the job. At the door, I announced that I could never scan another book and that, no, I would not be back Monday. By the time I quit, I had already scanned 23,712 pages, in 64 hours, an average speed of roughly 6 pages per minute, faster than many two- or three-month veterans. A stack of all those pages would stand 9.8 feet tall. I remember writing dramatic passages in my notebook about “the bone-crushing factory work of the new information economy” during my short tenure, but looking back through my notes, such observations—indeed, any mention of the Internet Archive—are strangely absent.