

*Since the material in this column often represents personal views tempered or sometimes weakened by memory, the editor invites other opinions and evidence.*

*Editor's note: Shane Greenstein relates an anecdote centered on the nature of historical research in computing. It is valuable both for what it says about the current state of historical awareness in our field, and for its insights on conducting research in uncharted territory.*

### The Tape Story Tapestry: Historical Research With Inaccessible Digital Information Technologies

We all recognize that digital technology has lowered the cost of storing data and undoubtedly helped produce an increase in the amount of research data available. Yet, the invention of digital technologies alone was not sufficient to make the data useful. In the first place, stored data has to be inherently "useful" for answering a worthwhile research question. Second, just because information is "less costly" to store does not necessarily make it more "accessible," especially for future research efforts.

This essay has two simple observations, neither of which is widely appreciated. First, if machine-readable data is to be made accessible to future generations, then a number of complementary activities must be performed, many of which we all take for granted when using other storage mediums, but do not exist for machine-readable data. Second, digital technology makes it incredibly easy to destroy what could not have been gathered without its invention in the first place. While digital technology makes the storage of more information less costly, it also makes the misplacing, erasure and elimination of historical data less costly.

The result is that much useful information stored in machine-readable form is becoming lost to future generations of researchers. Often when the original purpose of private data is fulfilled, the user has little incentive to protect the data for historical purposes. When data is placed in the public domain, if it is placed in the public domain at all, it is not organized with the appropriate complementary aids. Hence, despite all the formal mechanisms designed to help researchers retrieve original archival sources, researchers still find that unsystematic informal communication is still essential for retrieval of digital information that is not kept with historical interests in mind.

I personally experienced the consequences of these problems in the spring of 1987 and was fortunate enough to succeed despite them. In August of 1987 I received thirteen computer tapes in the mail. These were the last

remaining copies of data recorded between 1971 and 1983.

Describing the hunt for these tapes is not a trivial story because the data has great research value. The story contains a number of improbable events which lead to success—including personal contacts, chance encounters, near misses, 11th-hour reprieves, and a maze of federal government agencies. Like any decent mystery, this story also has its share of heroines and heroes, many blind alleys and red herrings, and an ending full of lessons for the wise.

Yet, in the end, it is those lessons that make the story worth the telling. The tapes were obtained not through the use of a special index, not with the aid of some of the best reference librarians on the Stanford University campus, and not as a result of a formal records management system. All those mechanisms failed because the information never was properly stored for future generations, or more to the point, the historical value of the data was never recognized. The required documents were obtained because I was fortunate enough to contact a number of the people associated with developing and using that information and who were kind enough to help with my search.

#### *The Unwinding Of The Tape Story*

Since the late 1950s, when it was a relatively easy task, up until the present, when it became nearly impossible, one Federal Government agency or another has tracked the state of the entire government's computer equipment inventory each year. The reasons they did this, and consequently, the form of the collection changed over time, but the basic task did not. Each year every federal agency and department completed a survey designed to record and categorize their holdings of computer systems. The oversight agency tabulated and published a summary of these answers. The sum of these surveys constitutes a federal inventory of Automatic Data Processing (ADP) equipment. Each year's inventory is complete, carefully assembled, and informative.

The inventories provide an unprecedented record of the federal government's involvement with information technologies. For example, they document the growth of total federal government holdings from 531 "processors" in 1960 to 18,474 in 1982. The inventories show the significance of the government's early involvement in purchasing computer equipment when the industry was young and technologically immature. To the knowledgeable observer, they provide a record of system configurations that track the changing technological norms of computer equipment use in the country. Later inventories contain evidence of the commonly held perception that the government's systems have aged, heavily burdened by older investments that were not replaced as rapidly as in private firms.

In other words, a collection of these inventories pro-

