THE WALL STREET JOURNAL.

http://www.wsj.com/articles/big-data-before-the-web-1450047074

ARTS | BOOKS | BOOKSHELF

Big Data Before the Web

In the 1950s, social scientists tried to preserve an archive of human experience on microcards. Their experience is a parable for our digital era.
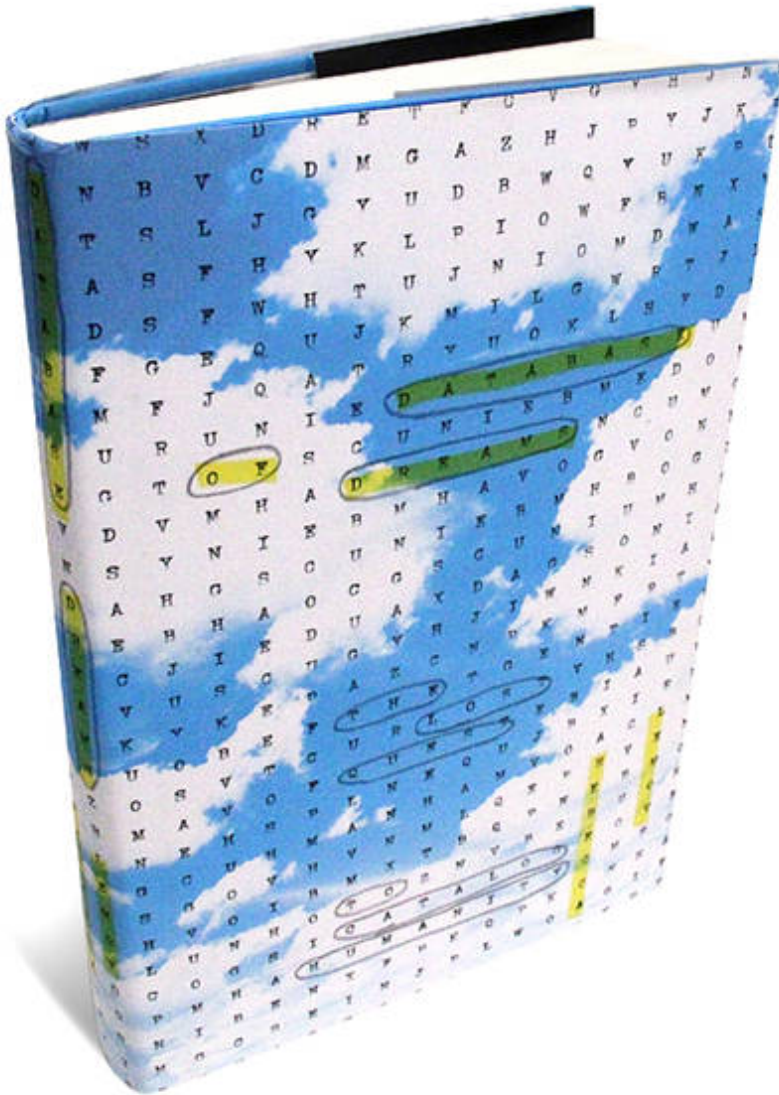
By **EVAN HEPLER-SMITH**

Dec. 13, 2015 5:51 p.m. ET

Sometime in the early 1950s, on a reservation in Wisconsin, a Menominee Indian man looked at an ink blot. An anthropologist recorded the man's reaction according to a standard Rorschach-test protocol. The researcher submitted a copy of these notes to an enormous cache of records collected over the course of decades by American social scientists working among various "societies 'other than our own.' " This entire collection of social-scientific data was photographed and printed in arrays of microscopic images on 3-by-5-inch cards. Sets of these cards were shipped to research libraries around the world. They gathered dust.

In the results of this Rorschach test, the anthropologist saw evidence of a culture eroded by modernity. Sixty years later, these documents also testify to the aspirations and fate of the social-scientific project for which they were generated. Deep within this forgotten Ozymandian card file sits the Menominee man's reaction to Rorschach card VI: "It is like a dead planet. It seems to tell the story of a people once great who have lost . . . like something happened. All that's left is the symbol."

In "Database of Dreams: The Lost Quest to Catalog Humanity," Rebecca Lemov delves into the ambitious efforts of mid-20th-century social scientists to build a "capacious and reliable science of the varieties of the human being" by generating an archive of human experience through interviews and tests and by storing the information on the high-tech media of the day.

For these psychologists and anthropologists, the key to a universal human science lay in studying members of cultures in transition between traditional and modern ways of life and in rendering their individuality as data. Interweaving stories of social scientists,

Native American research subjects and information technologies, Ms. Lemov presents a compelling account of "what 'humanness' came to mean in an age of rapid change in technological and social conditions." Ms. Lemov, an associate professor of the history of science at Harvard University, follows two contrasting threads through a story that she calls "a parable for our time." She shows, first, how collecting data about human experience shapes human experience and, second, how a high-tech data repository of the 1950s became, as she puts it, a "data ruin."

During the early 1950s, a young psychologist named Bert Kaplan began to worry about the fate of what he and fellow social scientists called

DATABASE OF DREAMS

By Rebecca Lemov
Yale, 354 pages, $35

"human documents"—the Rorschach-test results, along with notes on stories and dreams. Fresh from fieldwork among the Zuni in New Mexico, Kaplan had firsthand experience of the difficulty of gathering personality data. What was more, he felt that such data was much more important than most of his colleagues realized. Examining his own data, he had not found a typical "Zuni personality" but rather divergence and individuality, bucking the scholarly consensus that personality was shaped by culture.

Yet after they had written their articles and monographs, his fellow social scientists typically set their data aside to be lost or buried. As Ms. Lemov puts it, "it was not only 'far-off' ways of life that were disappearing in the wake of modern changes, but also the

data that documented this process were themselves in danger of disappearing." Preserve the culture, Kaplan believed, but also preserve the data.

He dedicated himself to building such a collection, persuading colleagues far and wide to send him copies of their notes and test results. Among the various contributors to this effort, Ms. Lemov paints a particularly vivid portrait of Dorothy Eggan, a woman without a college degree who became a leading figure in Hopi anthropology based on her work turning dreams into data. To store and reproduce the collection, Kaplan settled on an easy-to-use format that had recently been developed by a librarian-inventor: the microcard.

Ms. Lemov's book "was supposed to be about dreams" collected from the Hopi and other research subjects, she writes, but it also "ended up being about technology." Tracing the genealogy of Kaplan's chosen storage medium, Ms. Lemov tells a fascinating story of the development of microphotography. This assemblage of tools for shrinking, storing and reconstituting images began as a Victorian drawing-room curiosity. By the time Kaplan seized upon it, microphotography had been used in various ways: to record information that pigeons could carry in secret messages; to enable correspondence between soldiers and the home front; and to store data for corporations. In all of these cases (as in more recent innovative formats for data storage—remember the floppy disk?), microphotography at once rendered information accessible and, through technological obsolescence, potentially inaccessible. On microform media, "compressed and miniaturized, documents were both secure and in jeopardy."

Ms. Lemov has a keen eye for such paradoxes. The secrets of Zuni experience "sit in field notes, data sets, and eventually a data archive even as they apparently evaporate from Zuni daily life itself." In Kaplan's project to catalog humanity, "the closer one comes or believes one has come to perfect knowledge, the more fragile, dream-like, and perilous does that summmum become." Ms. Lemov writes with a literary sensibility; T.S. Eliot, James Joyce, Wallace Stevens, David Foster Wallace and Stephen King all make fleeting appearances. The act of recording experience in writing, Ms. Lemov shows, could also shape it: Hopi dreams begin to feature cameo appearances by Dorothy Eggan and her dream notebooks.

Shifts in social-science methodology and the eclipse of microcards as a data-storage medium relegated Kaplan's catalog to "an early if not complete obsolescence." Exploring its significance today, Ms. Lemov writes, "felt akin to thrift store history, finding value in discards others have ceased to value." Ironically, Kaplan himself had been up to much the same thing in soliciting his colleagues' old data sets, unaware that

his own collection would itself fall into obscurity. Might other totalizing data-collection projects suffer the same fate? This is one sense in which "Database of Dreams" is "a parable for our time." It is a memento mori for Big Data.

Mr. Hepler-Smith is a Ph.D. candidate in the history of science at Princeton.