

Hidden pasts uncovered in Times' series

Data base analysis, 'shoe leather reporting' key in 6 month investigation by reporters

Robert Jackson
MICAR

In Florida, it's easy for people with criminal records to hide their past.

That's what *St. Petersburg Times*' reporters David Barstow, Susan Taylor Martin, Chuck Murphy, Bob Port, and editor Richard Bockman discovered after a six month long investigation into the abuse of a state statute which allows people charged with crimes to have their records sealed.

What the reporters found was astonishing. The statute created to give first-time marijuana offenders a second chance, was now giving offenders charged with crimes ranging from fraud to murder a second chance and sometimes a third.

The list of abuses was lengthy and required the four reporters to be creative in compiling the data, according to Port, who did most of the computer analysis. It was also the *Times*' first crack at computer-assisted reporting he said.

See "Hiding" on page 2



Bob Port is a general assignment reporter for the St. Petersburg Times. He was responsible for most of the data analysis during the six-month-long investigation, which culminated in the five part series.

Mainframe versus PCs: it's a matter of time, cost

Lisa Touye'
MICAR

As the number of papers doing computer-assisted reporting increases, so does the number of people searching for a system that meets their newspaper's needs.

Most papers use tape-drive systems, like 9-Track Express, to download tapes while some use their newspaper's mainframe.

The *Sacramento Bee* is one of the latter.

George Schlukbier, computer coordinator for editorial and library director at the *Bee*, uses the paper's mainframe to download information

for their projects.

Using a VAX network hook up, Schlukbier downloads the whole tape into the 2.4 gigabyte hard drive attached to his MacIntosh FX. The VAX network hookup allows his Mac to import records as quickly as the mainframe downloads them.

Then he uses FoxBase Mac to extract information from the database now on his hard drive.

The mainframe system helps keep the *Bee* competitive with the other papers in the area doing computer-assisted reporting, like the *Los Angeles Times*, *San Francisco Chronicle* and *San Jose*

See "Bee" on page 4

"It (the investigation) was a good example of shoe leather reporting combined with information analysis on a computer data base," Port said.

First the *Times* had to obtain the state arrest records and occupational data on magnetic tape, Port said. The *Times* hit a stone wall in their negotiations with the Florida Department of Law Enforcement, which happens to be "home" to the state's Crime Information Computer. The newspaper requested that the FDLE provide them with the state's arrest records on computer tape.

Florida is unusual in that certain information on the arrest records is considered public record, according to Port. "We requested the records on tape under the stipulation that all of the identifiers — names, addresses, date of birth — be removed from the record."

The FDLE denied the newspaper's request claiming, that while the arrest records — on paper — were considered public record, the tapes were not. Fortunately for the *Times*, records on magnetic tape falls under the umbrella of Florida's "sunshine law."

"What it boils down to is they were afraid that we, being crafty reporters, could use other records to get a match on the names," Port said.

The *Times* finally had to threaten a lawsuit to get the FDLE to honor the request, and although the FDLE finally agreed to provide the two tapes, they made the experience costly — after paying \$25,000 in legal fees, the *Times* spent an additional \$5,000 to obtain the tapes.

"That's a good indication how serious the *St. Petersburg Times* was in obtaining that information," Port said.

Once the *Times* had obtained the tapes, Port imported the information into a PC with two 330 megabyte drives using 9-Track Express. He was then able to manipulate the data using XDB and SQL.

Port explained that the FDLE tapes were strictly used for statistical analysis. "For instance, we were able to use the information on the tapes to show which counties had the most requests for sealed records," he said.

Ironically, every person who was named in the

series as having a sealed record was identified means other than from the arrest records.

Port built a data base of 3,000 people who had gotten records sealed. Other public records were then used to find out more about the individuals. "It was piecing together a big jigsaw puzzle," he said.

Building the data base required reporters sit through scores of court hearings where defendants requested that their records be sealed. Then reporters obtained the defendants police and court records before a judge ordered them to be sealed.

Reporters also spent hours in the *Times*' library, where cases reported by the newspaper were checked against current court and arrest records to see if the record had been sealed.

"What it boils down to is they were afraid that we, being crafty reporters, could use other records to get a match on the names."

-- Bob Port

"I guess you could say that we were crafty reporters," Port said.

The various techniques turned up plenty of abuses, and the *Times*' five part series entitled "Hiding the Past," may cause Florida lawmakers to repeal the statute, according to Port.

"We expect the law to change as a result of the series," Port said. "Legislators are already talking about either repealing or reforming the law."

The cost of the endeavor has not turned the *Times*' top management sour on computer assisted reporting, according to Port.

"Most of the editors at the *St. Petersburg Times* would say that they don't have a clue about computer assisted reporting," he said. "But the power of the computer has started to kick in and they are becoming accustomed to the concept."

Editors note: Some information in this report was reprinted by permission of the *St. Petersburg Times*.

Paper chase:

Mainframe, helpful librarian solve paper woes for Post-Tribune reporter

Robert Jackson

MICAR

Gary, Indiana *Post-Tribune* reporter Joseph Conn knows his city's census data — intimately.

That's because Conn spent the better part of 20 hours pouring through reams of the data, trying to figure out how many of Gary's west side residents would be displaced if the city received a multi-billion federal construction project to expand its airport.

"According to airport boosters the expansion could create about 140,000 jobs and bring billions of dollars into the area," Conn said.

So Conn began with the 1991 Census data. He began pulling block numbers off paper maps with the idea of comparing that information with the proposed "footprint" of airport expansion.

"I had already invested about 20 hours in the task and had not even begun to sort the data and marry it to the list of block numbers," Conn said. "It was like using a spoon to clean a stable."

Conn figured it would take another 20 to 40 hours just to come up with some useable information. That is until he learned about a redistricting computer program called GEO District, and a librarian named Steve Fisher.

"I learned that the state had purchased this program, which is similar in some respects to Atlas GIS," Conn said. "There are six terminals on the state's GEO District system but only one allowed public access."

So he made an appointment and drove to Indianapolis. That's where he met Fisher.

"Basically, I navigated and (Fisher) steered," Conn explained. "I gave directions to him from a street map with the proposed airport 'footprint' drawn on it, and he used a mouse to trace the footprint onto the computer screen."

Using the program's lasso and "point-and-shoot" tools, Fisher placed the footprint on a map of U.S. Census Bureau blocks for the area specified by Conn.

Once the map was drawn and all footprint blocks were selected and filled in with color by the mouse, Fisher

querried the data base for households, total population and population by racial groups within the specified area.

What Conn found amazing was that it took very little time.

"(the data base query) only took five minutes, and copying the reports from that query onto a 3.5-inch diskette took about five minutes more," Conn said. "Making the map took roughly one hour."

He was even able to verify the GEO District-generated block numbers by checking them against the paper sheets he had already developed.

"I found about 10 block numbers that the program didn't pick up along the footprint margins," he said. "It seems the program had difficulty picking up some of the smaller census blocks or we failed to see that certain blocks were not selected or colored in."

"On the flip side, the program did pick up about six blocks that I'd missed when doing the listing by pen and paper, so one effort aided the other," he added.

This newsletter is a publication of the



**Missouri Institute for
Computer-Assisted
Reporting**

120 Neff Hall

University of Missouri

Columbia, MO 65211

*MICAR is interested in attaining any
information, ideas or stories related to
computer-assisted reporting for future
issues of UPLINK*

*If you wish to contribute, please mail
your story or idea to the above address. Or,
call us for a fax number.*

-- Robert Jackson, Editor

Mercury News, Schlukbier said.

When the Census released the Public Law data files earlier this year, the *Bee* was ready with a survey of minority representation on city and county governments, Schlukbier said.

When the *Bee* got the tape, they checked representation with minority population figures from 1990 and 1980 and had a story out by deadline that showed serious discrepancies in population and representation.

For example, one area in Los Angeles that was 92 percent black had no black representation on city council, Schlukbier said. Analysis also showed that Hispanics were better represented on city and county governments than blacks, said David Jensen, special projects editor at the *Bee*.

"It's almost impossible to find trends or anomalies by looking at the Census' STF1 file print out," Jensen said. "Those things jump out at you when you use the computer."

Schlukbier said he went with the mainframe system because it met his paper's needs which are:

- constant access to all the information in their 9-track tape library from any terminal in the paper.

- a system that allowed special projects reporters, graphic artists and editors to access and use information from the same database without passing around hard copies or diskettes containing the information.

- compatibility with the art department's MacIntoshes and their

Illustrator and DeltaGraph software.

"I decided from day one not to go with 9-Track Express," Schlukbier said. "I wasn't convinced that loading and reloading 9-track tapes each time I needed different information was the way to go."

Schlukbier said it takes 15 minutes to download a tape with variable length records and packed fields into FoxBase using the mainframe.

He said that he found tape drive hardware for the mainframe to be more reliable and user friendly than those for stand alone personal computer systems.

"I wasn't convinced that loading and reloading 9-track tapes each time I needed different information, was the way to go."

-- George Schlukbier

The keys to the newspaper's system are the mainframe and the 2.4 gigabyte hard drive, which allows Schlukbier to download a whole tape with no problem. But the cost of a hard drive alone with that capacity may be prohibitive for most newspapers. The one Schlukbier uses runs about \$8,000.

Schlukbier said the site license for the VAX network hookup cost \$1,000 and \$300 for each Mac in the network. The database

program he uses, FoxBase, costs \$400.

The editorial department has an agreement with the computing department that when a tape is acquired it can be loaded immediately. According to Schlukbier, a reporter doesn't come in with a hot tape needing to get some information off it immediately for a story. Acquiring a tape can, and usually is, a lengthy process, he said.

Schlukbier works with a reporter when he downloads the tapes and conducts queries.

"There aren't hot tapes," Schlukbier said. "We do this for long term special projects. If we need quick access to information we go online or to our own databases."

Many newspapers do not have a mainframe that they can use to quickly download tapes.

That's where stand alone systems come in. Systems like 9-Track Express were created to make it easier for reporters to do their own computer analyses on something familiar to them — the newsroom PC, said Dan Woods, co-author of 9-Track Express.

"Many of the dinosaurs of the mainframe world have a siege mentality when the fact of the matter is that the future of reporting or any other analysis is on desktop," Woods said.

"Analysis is an iterative process by which answers to certain questions lead to other questions," Woods said. "If the reporter is not involved in that process he's not really reporting."