UPLINK
THE FORUM FOR COMPUTER-ASSISTED REPORTING
April 1992

Courant creates profile of murders

Grant Houston
Hartford Courant

After doing computer-assisted reporting for a few years, I find myself and my colleagues increasingly attracted to those phrases such as "a computer-assisted review of 1.2 million records found..."

It's imposing, comprehensive and satisfying to see that number in print when weeks and months of your life prior to computers was spent sifting and stitching together a few hundred paper files for investigative projects.

But recently I had a sharp reminder of how some simple nuts and bolts computer-assisted reporting — that uses only a few hundred records — can serve as a tipster, a backgrounder and an investigator, and end up having high impact.

Last October, Courant police reporters noticed a streak of unsolved murders involving minority [illegible] the Hartford area, and they mentioned it in a story in the back pages of the newspaper. However, the reporters had trouble going further or pinning down police investigators because they had no systematic files on the murders and our clip library is not electronic.

I decided to create a computer file for the reporters that would include all of the unsolved murders of minority women in the state during the last few years. I created one file in an XDB database from our clips that included details of the murders and demographics of the victims. I knew it was incomplete, but it gave us about two dozen solid cases.

With a little sorting and comparing and working with the police reporters and Leah Segal, our Courant researcher, we culled the list to about six or seven cases. The reporters then began questioning police in the Hartford area, using the details from the database. The police thought we had a tipster and conceded that they had quietly been looking at links among five murders. That broke the first story that there might be a serial killer on the loose.

The police reporters, Gina Brisgone and Andrew Julien, and I then collected more information for the database by visiting the sites where the victims were last seen and where their bodies were dumped and by interviewing victims' friends and families.

Then I went to the state's medical examiner and asked for his computerized files. The medical examiner was worried about confidentiality but gave us several fields of information.

Unfortunately, he didn't know how to download the information to diskette. I asked him to just give me

See "Murder" on p. 3

MO Education Department releases test scores to MICAR

Russ Buettner
MICAR

The Missouri Department of Education wrongfully imprisoned valuable data, at least until continued [illegible]ment threatened to require effort.

Since beginning an education project last September, the Missouri Institute for Computer-Assisted Reporting has acquired various computer files related to education. The Missouri Mastery and Achievement Test, this state's version of a standardized test, seemed a logical addition.

The department's usually helpful programmer could tell me nothing about the file. So I called John Jones, director of school data. In his lofty lobe, journalism serves no purpose beyond nuisance.

"Russ, I don't keep that information, and I don't want it," he said. "If I had it, every newspaper in the state would be calling me for their little score sheets."

You can imagine the sickening image: a director of data...providing data.

See "Test" on p. 4

Despite XDB's popularity with computer-assisted programs across the country, some rumblings of dissatisfaction have surfaced in several newsrooms. The problems cited may have existed in earlier versions of the software, but with some reporters switching to other software, the question now is . . .

Will early version woes hurt XDB?

David Raziq
MICAR

For two weeks in July 1989, Bob Paynter was supposed to be on vacation. But during his stay at a North Carolina beach, he frequently found himself on the phone. The problem: his software. Paynter specializes in computer-assisted reporting for the* Akron Beacon-Journal, *and he says his database program of choice, XDB, was giving him major problems.

"I was very concerned," says Paynter. "I was involved in a project involving campaign finances that required me to move a large database back and forth between XDB and Dbase."

Paynter says his troubles began when he tried to import the information back into XDB. The program wouldn't accept it. So he placed the first of many calls to the software company.

"There were people joking in our department that suggested I demand a development fee from XDB," Paynter says.

In the last few months the staff of MICAR has heard a number of complaints from XDB users. Bob Paynter's story is representative.

For example, Paynter himself finally ended up shipping his entire database to XDB. During his break, XDB's technical support eventually solved his difficulty with a new import engine.

It wasn't long before he experienced another set of problems — indexing a large file.

"It would take a split second to index a file of almost two million records so I knew something was screwed up," he says.

When Paynter queried about the values within the file, his computer came up blank.

So after another month trouble-shooting with XDB's technical support staff, Paynter got the file to index. But he found a new bug: Given the exact same query, the program would return different answers .

And that, he explained, was the last straw.

"I said 'I just can't trust this,'" he says."And I talked to other database managers, and nobody had ever heard of XDB." So after consulting friends, Paynter switched to the Paradox database system.

Ed Foldessy of the **Wall Street Journal** has also abandoned XDB for his record analysis. "XDB is not graceful with numbers," he says. "I encountered every conceivable problem working with it."

According to Foldessy, one of those problems included sorting fields. "It wasn't giving a good sort," he said. "You might have .01, .02, but then you'd have a number that was out of order."

Like Paynter, Foldessy says he "will never use XDB again," and instead relies on Dataease, his corporation's recommended program.

However, many reporters stand by the XDB product. Elliot Jaspin, the director of the Missouri Institute for Computer-Assisted Reporting, is one of the pioneers in this reporting technique. He has had a long standing relationship with XDB.

"When I was teaching at Columbia University, I began searching for a database program," says Jaspin. "So I went down to **PC Magazine**'s software library and tried them all. And XDB was the winner, hands down."

Jaspin lists speed, flexibility and ease of use as some of the product's strong points. Because he also uses XDB in his computer-assisted reporting seminar, he says the program has some other advantages. "It's a good teaching tool," he says. "It has a simple interface, and it has good implementation of structured query language."

But Jaspin also readily admits that the program has had an "unacceptably high number" of problems. Specifically, Jaspin mentioned problems indexing large files, and he has noticed an "unevenness" in XDB's support staff.

According to Mike Waters, XDB's Director of Sales, most of the problems were found on an earlier version of the software.

For example, Waters explained that Paynter's indexing problem was due to the index pointer being defined as a small integer. "So once you got up to a little over 31,000 unique indexes, you ran out of numbers," he says.

Waters said his company corrected those types

See "XDB" on p. 4

Murder: from p. 1

z ...n printout, and we borrowed a scanner from our ...ertising department and scanned about 200 records into the computer. Then I added some information from a death certificate tape we had recently acquired.

As the reporters were using the database to help do individual stories on the victims, I linked the various files together and started sorting and comparing while consulting experts on serial murders. It appeared there were similar murders throughout the state. On one interesting run, it appeared the murders generally occurred on a quarterly basis.

We also talked to a university professor who had data on 16 years of homicides in Connecticut. At our request, he did a quick run with our criteria and found the recent series of unsolved murders was a definite — if not statistically proven — blip on a chart of such murders.

A few interviews later and we had two more stories that expanded the possible range of the murders and showed that police in different jurisdictions were investigating similar murders without talking to each other.

By now we had some investigators who were sure we had access to confidential files and wanted to know what we knew, and some of whom were full of culpas. One investigator told me point blank that we had done what police should have done months before.

As a result, the state's chief prosecutor convened a task force in January composed of police departments to look into at least 11 unsolved murders.

A caveat: We never asserted there was a serial murderer at work. We simply established a pattern of murders that the police agreed should be seriously investigated. Even if two or three killers are not responsible for all the murders, would it be good news that there were 11 killers instead?

We also focused attention on the murders of minority women — some of whom were prostitutes — that had not been there before.

Sadly, many cities have a collection of unsolved homicides of black or Hispanic prostitutes that are dismissed as routine (murder is supposed to be an occupational hazard). In several other states, serial murders have plied their trade for years because society generally could care less about those kinds of victims.

And our stories also allowed the newspaper to reaffirm its ties to our inner-city neighborhoods. We were thanked many times by victims' families and friends who thought these were the kinds of cases that the Courant wouldn't care about, much less investigate. ■

• • • • • • • • • • • •

***Brant Houston is an investigative reporter who runs the computer-assisted reporting program at* The Hartford Courant.**

Bytes, bits and nibbles...

Racist cops and lethal working conditions were a few of the problems exposed by the winners in this year's IRE reporting competition.

Mike Casey and **Russell Carollo** of the **Dayton Daily News** won for their series "Worker Safety in America." The team used computer analysis of records from the Occupational Safety and Health Administration to document the failure of that agency to protect American workers.

Carolyn Tuft of the **Belleville (Ill.) News-Democrat** won for a four-month computer-assisted investigation of city traffic tickets that revealed the city's police department was targeting black motorists in ...attempt to deter them from entering the ...ity.

Winners will receive their awards in June at the IRE National Conference in Portland, Ore. ■

This newsletter is a publication of the

Missouri Institute for Computer-Assisted Reporting
120 Neff Hall
University of Missouri
Columbia, MO 65211
(314) 882-0684

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

Test: from p. 1

I asked him who had the data. He said the department contracted with the Educational Assessment Center, another state agency, to administer, score and store the test results.

So I call Sharon Shattgin, assistant director of the center.

"See technically, we don't own the data," she told me. "Technically, it's the property of each school district. We contract with each superintendent. The data is public information, but you have to get a release from the superintendent."

But I didn't want one district. I wanted all 453 districts.

"I guess you'd have to get a signed release from each district," she said.

Even over the telephone, you can tell when someone is talking through a smirk. I asked her if there was anyone else I could speak to about this.

"Have you spoken to John Jones?"

As ridiculous as it sounds in itself, this arrangement, clearly designed to block access to the data, had a much worse effect. As a result of not "owning" the data, the department of education has never analyzed test results in their entirety, nor by gender, race, rural versus urban, or in any other way. It only publishes an annual mean based on a 10 percent sample.

In other words, the state spends millions of dollars administering the tests, but does not use it to uncover weak spots.

Although the leaders in Missouri education seemed to think they had cleverly locked up the data, even from themselves, they had unknowingly propped open a window for us.

XDB: from p. 2

of problems with the 2.4 version of XDB and currently markets a 2.41 version of the software.

Waters also explained that the once-limited indexing function of the 2.3 software could have caused Paynter to get different answers for the same query. "When you do the query, the optimizer inside is actually building indexes on a slide," he said. "So if he was dealing with a lot of records with the 2.3 version, he could have gotten inconsistent results if he got more than 31,000 on an index scheme."

However, Waters said he was unaware of any complaints about sorting fields of numbers, as was the problem Ed Foldessy experienced.

"As a matter of fact," he says, "I offer free products to any of these guys that made a comment to you. I'd love for them to try them again." ■

Jones, incarcerator of data, had provided us earlier with a computer file containing the title, name and mailing address of every school administrator in the state. It took less than an hour to download the file, move it into an XDB table and export the table as a WordPerfect mail merge file. I then wrote a form letter to the superintendents explaining this ridiculous situation and informing them of their responsibility under the state's Sunshine Laws to respond within three working days.

But I never intended to actually print, let alone mail, those letters.

Instead, I wrote a letter to Otis Baker, the director of elementary and secondary education, explaining our efforts and the apparent necessity of the mailing. I told him our form letters and return envelopes were addressed directly to him "to insure their authenticity," and that each letter referenced his name and phone number "should any of the 453 superintendents have any questions regarding the necessity of this procedure."

I told him we had scheduled everyone on our staff for an envelope-stuffing party on March 3, three working days after he received the letter. If I didn't hear from him by then, out the letters went.

When I arrived at the MICAR office on March 3, I found that Mr. Baker had called.

"He said he needs to talk to you TODAY," said MICAR Administrative Assistant Michele Ward. "He said he just got your letter today. He'll be in meetings all day, but his secretary will interrupt when you call. He definitely needs to talk to you today."

Otherwise deaf bureaucrats seem to perk up to the threatening sound of 400-plus letters and phone calls thundering in over a few days.

"I just got your letter today," he said when I called. "Don't send out those letters. You've gone to too much trouble already. I'll send out a release letter today."

In an eventual story, we might tell Missouri's citizens more than they know about their education system. We might also tell Missouri's education leaders more than they apparently want to know about the jobs they do. Yet I doubt the Department of Education will begin graciously providing files.

Nonetheless, maybe computers can help us free more wrongfully imprisoned data by presenting the jail keeper's greatest menace -- the threat of imminent labor. ■

• • • • • • • • • • • • •

Russ Buettner is MICAR's research assistant and a graduate student at the University of Missouri School of Journalism.