UPLINK
THE FORUM FOR COMPUTER-ASSISTED REPORTING
v12
October, 1991

Newsroom conflict:

Anything can (and possibly will) go wrong when too many DPs work with your data

Jerry Uhrhammer
Tacoma Morning News Tribune

I felt vaguely uneasy when I attended MICAR last May and listened to Elliot Jaspin talk about the importance of maintaining newsroom control over computer-assisted reporting projects.

Before leaving for Missouri, I'd been talking to our newsroom computer systems technician about the hardware needed for a computer-assisted reporting program.

He didn't think that a separate nine-track tape drive was necessary. He was convinced it would make better sense — and be less expensive — to download nine-track government tapes on the newspaper's VAX business computer, then transfer the data to the newsroom network's servers where it could be accessed with a PC.

See "Conflict" on p. 3

TRI working guide available at MICAR

The Environmental Protection Agency's Toxic Release Inventory (TRI) has been one of the most popular databases used by the readers of UPLINK. It has been the source of excellent industrial pollution stories in Florida, California, Connecticut, Ohio and elsewhere.

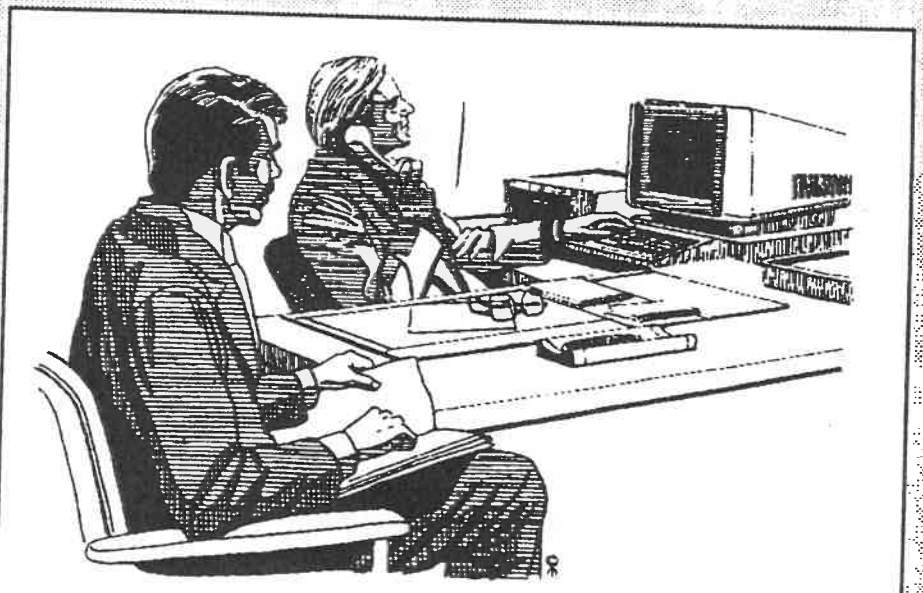
But many reporters still have not used the database.

In June, N.T.I.S. made the 1989 TRI available on computer tape. We have found a way to make

See "TRI" on p. 4

'Daily News' links up with county

Info access 'just a few keystrokes' away



Jim Ripley
Dayton Daily News

A personal computer a few feet away from the metro desk is forever changing the way reporters research county records at the *Dayton (Ohio) Daily News*.

The PC provides a direct line into the county's computer. With a few keystrokes, reporters can examine current Montgomery County property tax records, deeds, mortgages, auto titles, civil court dockets and file indices, and divorce and domestic violence fil-

See "Dayton" on p. 2

ings. In coming weeks we expect togo online with the county jail and criminal court, and we're negotiating for access to the board of election's voter registration files.

Making the technological link for the system was far easier than getting past county officeholders.

We made a formal request for access last September. The officeholders turned it over to an attorney to examine privacy issues and the applicability of Ohio's Public Records Law. The attorney decided the law was on our side.

The Dayton Daily News is paying Montgomery County $125 per month to service and maintain the hookup. There is no hourly or per minute charge. We also have to pay $480 per year for a dedicated telephone line.

The equipment, including a special modem and a device to enable our PC to emulate a county dedicated terminal, cost $1,180.

One drawback to the system is that we can't download data onto a diskette or hard disk drive. However, we can print copies of the records.

The newsroom's computer link to the county showed its potential the same day it was connected.

Reporters were responding to tips that a financially troubled department store chain might close its downtown store. With an instruction manual at his side, deputy metro editor John Erickson typed in the codes that enabled us to gain entry into property tax records. Within minutes he determined that the store was $800,000 behind on its taxes.

For courts reporter Rob Modic the hook-up meant a page one story. Modic had a letter from a college student who believed her immigration attorney had skipped town with her money. He also had a motion filed by the local bar association seeking to reassign the immigration attorney's clients.

An hour on the system with Erickson convinced Modic he had more.

— The county's electronic records showed the attorney had purchased a $305,000 home with a $275,000 mortgage less than a year ago.

— They showed that he had taken out a second mortgage for $40,000 nine days before leaving town in July.

— They showed he had bought a $40,000 Mercedes.

— They showed federal tax liens of $36,825 had been placed on the man's home and $22,504 on his business.

— They showed a list of creditors who had sued him along with attorneys' names and telephone numbers.

The county's computer database only summarized court records. Modic still needed to read some of the files for details. The newsroom hook-up helped him determine which cases held the most promise. Within minutes of closing time at the courthouse, he telephoned clerks, gave them the case numbers, asked them to pull the files and told them he was on his way.

Reporter Rob Modic used the county link-up to file a page one story.

By 7:30 p.m., Modic moved a page one story that reported, in fascinating detail, the shell game this attorney had been playing.

Modic estimated it would have taken him at least another day to turn his story had not a few computer commands taken the place of a maze of offices staffed by bureaucrats who hate being bothered.

More than 40 reporters and editors, including feature and sports writers, have been trained on the system.

One of them is minority affairs reporter Edwina Blackwell Clark. Clark used the PC to tap into county auto title records. She was working a deadline story about a man accused by authorities of being a broker of undocumented workers. The county computer divulged that the man bought seven cars last year and owned two apartments, which authorities later said were for the undocumented workers.

Jim Ripley is the metro editor for the Dayton Daily News.

Conflict: from page 1

That's why I was uneasy. Jaspin was telling us hat if you're going to use a computer to analyze government data from nine-track tapes, that's a reporting function. And it's important that the reporter have control of the software and hardware involved in the process.

I happen to be a computer enthusiast as well as an investigative reporter. I also know my own limitations — I'm a non-programmer and non-technician. And I had no illusions about being given access to the business computer to load my own tapes. I knew that if we used the VAX system, I effectively would lose a good deal of control over what I was doing.

As it turned out, Jaspin's words foreshadowed a summer of frustration and delays that effectively stalled my newspaper's venture into computer-assisted reporting for months.

Returning from Missouri, I wrote a memo that raised the control issue. But management opted instead to first try the VAX-fileserver-network approach. Why spend $10,000 on new hardware if you can do the job as well, if not better, with existing resources?

I soon learned first-hand what Jaspin had been talking about.

Just the simple act of loading a nine-track tape involved working through two tiers of bureaucracy and across departmental lines. First I had to give the tapes to the newsroom computer technician, who then took them to the data processing department. The data processing professionals (DPPs) actually loaded the tapes on the VAX.

The technicians and DPPs are stretched thin with their own projects and I learned that an assurance that something would be done the next day or the next week was not something to bank on. Days had a way of stretching into weeks.

When things did get done, the results were not always predictable.

The first tape loaded contained 1990 census data. But when I tried to extract data into my XDB database program, my 386sx computer labored mightily for a long time and brought forth absolutely nothing. I tried again and again. Still nothing.

Thus began a recurring pattern. I was the only one who knew about my XDB software. And when something didn't work, it was up to me to find out why. I spent hours on the telephone calling people familiar with XDB and census tapes.

Elliot Jaspin provided the clue that solved the first mystery. The VAX computer did not recognize the character/linefeed signal that marked the end of a record. The absence of such a command left the computer huffing and puffing through 170 megabytes of census data without finding the end of the first record.

When the VAX reprogrammed and the census tapes were loaded again, more funny things happened.

Asked for specific information on a certain census tract, the computer began spewing out data of all shapes and sizes.

I finally realized I hadn't set up the right filters to narrow down my request. But when I did, I still got two "hits" where I should have gotten only one.

After much detective work, I determined that the tapes had been loaded with an incorrect record length.

The good news was that all we had to do was reload the tapes using the correct record length. The

See "Troubles" on p.4

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

Troubles: from page 3

bad news was that data processing was busy with end-of-the-month business-side tasks. My tapes had to wait.

More bad news followed. After end-of-the-month time, the DPP who loaded the tapes left on vacation. Thirteen days elapsed before the tapes were reloaded.

In the interim, I learned another lesson about what can happen when you aren't in control of your own data.

Earlier, we had loaded a different governmental database into the fileserver. But when I finally examined the file, I found the data unreadable.

A couple days later, after getting advice on how to cure the problem, I tried to access the database again and found it missing altogether. I had been moved, without my knowledge, to make room for the anticipated reloading of the census tapes.

I was told the data would be restored the next day. Nine days later it still hadn't reappeared. I finally stopped asking about its whereabouts because I had new worries.

The 40-megabyte hard drive on my computer didn't have enough memory to handle the loads of data I was pulling into XDB. The solution: reprogram the system so that I would do all my work in the network's G-drive, which had plenty of memory.

That sounded like a pretty good idea, but when we tried it, my XDB software refused to work in G-drive.

XDB technical support finally figured out what was wrong. The AUTOEXEC batch file written to put XDB on the G-drive contained a command referring to C-drive, which was overriding all subsequent instructions to switch to G-drive.

The VAX-fileserver-networ approach turned out to be som thing less than advertised. For much of the summer, it seemed as if getting our computer-assisted reporting project off the ground was operating according to Murphy's Law: Whatever can go wrong will.

In retrospect, it illustrated again the basic truth of the adage, too many cooks can spoil the broth.

The troubles of summer had a happy ending. My editors agreed that we had to bring all of the program back inside the newsroom, giving us control over our own data.

We bought our own hardware for the newsroom, a 486-33 mhz computer with a 330-mg hard drive, a Qualstar 1260 tape driv and a printer. Now it's up to m a reporter, to make it all work. As it should be.

TRI: from page 1

using these tapes irresistible. MICAR has produced A Reporters Working Guide to the Toxic Release Inventory.

This short, step-by-step guide helps reporters download TRI into their computers, organize the data so that it is usable, and manipulate the data to produce meaningful reports

And because specific SQL commands are included, reporters using the guide will be able to:

— identify the major polluters in their area.

— list the quantities of toxic chemicals discharged by state, county or zip code.

— find out what toxic chemicals are used for.

— list the most common types of discharges.

The guide, which is really all you need if you already own the TRI database, costs $30 by itself. But for $50, MICAR will also include the NineTrack Express record profiles needed to download TRI.

If you do not own TRI, cannot afford it, or simply do not have a tape drive, for $350 MICAR will also supply TRI data concerning your entire state on 3.5" diskettes. These diskettes contain unabbreviated data, unlike that which N.T.I.S. sells.

Finally, if you wish to purchase the entire TRI database on computer tape, MICAR can supply it for $500. To order call MICAR at (314) 882-0684.

Bringing the classroom to the newsroom

By providing computer training to reporters, The News & Observer keeps edge on competitors

Lisa Touye'
MICAR

Most newspapers start and run their computer-assisted reporting program with one trained person; few papers train more than half of their staff. But that's what *The News and Observer* in Raleigh, North Carolina did.

"We made a decision early on that we would not restrict our equipment and knowledge. Access to databases would be made available to everyone interested in the newsroom, even interns." -- Pat Stith

"We made a decision early on that we would not restrict our equipment and knowledge," said Pat Stith, a reporter at the paper. "Access to databases would be made available to everyone interested in the newsroom, even interns."

Management knew that not everyone who was trained was going to become an expert, but it wanted to eliminate the mystery surrounding computers and databases.

"We attempt to instill in our staff that looking for and through databases is a routine part of the job," Stith said. "When you have a story idea you go to the library, interview sources and talk about it with your colleagues. Now you also look for databases with something additional that's important to the story, acquire the database, do the work on it and use the information."

To make databases a more routine part of the job management has held training programs since July of this year for anyone interested. Six people are trained in each class, which is held from 8 to 10 a.m. for two weeks.

"We start off as if they don't know anything...not even DOS," Stith said. "We teach them DOS, FoxPro, tapes, NineTrack Express, R&R (Report Writer) and MapInfo."

The classes are divided into two groups. One group loads the tapes, and the other writes reports, Stith said.

Reporters work on the paper's 17-station computer network, which is served by a Hewlett Packard 486. The paper has two additional 486s and one Qualstar tape drive with a total of 2.2 gigabytes of memory. One of the additional 486s is used to load tapes and do mapping, the other is a backup used for projects with large, complex calculations.

For some newspapers, starting a computer-assisted reporting program often meets with resistance from management or financial capabilities. But that didn't happen at *The News and Observer*.

"What's different here is that we were not required to prove that the concept of computer-assisted reporting works to our management, " Stith said. "It was accepted as a given. The general outlook is that this is an expected thing to do.

"Some newspapers, and this is an impression of mine, use database analysis on a few major stories," Stith said. "We intend to use it on any story where the time and expense of acquiring the database is negligible compared to what it adds to the story."

For example, the paper acquired the state's death records for the past 15 years on tape for $18.97 and used information for a story about cremation, he said.

See "Training" on p. 6

Training: from page 5

Along with building the network and teaching the system, *The News and Observer* has created a manual and a list of 3,000 databases available from the government. The list grows as reporters looking into a subject find what the state has available on it.

So far, the paper has acquired more than 25 databases—some federal, most state, Stith said. One agency in North Carolina keeps all the computer records for the state. The paper pays State Information Processing Services charges, which is the amount the state pays to get access to the computer records, but it won't pay more than the state does, Stith said. For example, a database of corporate companies and limited partnerships cost *The News Observer* a total of $43.

The public record law in North Carolina is particularly helpful because everything in the state is a public record unless another law specifically says it isn't, Stith said. When some agencies do deny the paper access to public records, *The News and Observer* goes to court.

"We have a well deserved reputation for not taking prisoners when people try to bend the law," Stith said. "We're polite and we say 'please,' 'thank you,' and 'may I,' but when we're denied access, we don't stand still or walk away We insist on making these agencies obey the law. We've done it for years with hard copies and now computer tapes."

When the paper decided to do a story about what happened when people moved away without a forwarding address and left safe deposit boxes, checking accounts, stock dividends, etc. behind, it ran into trouble getting the tapes it wanted from the state.

In North Carolina it's against the law for people who find this money to keep it—they must turn it in to the state who then must make an attempt to try to find the people and return it to them, Stith said. The state can't use the money; it must hold it so the people or their heirs can claim it.

When the paper first asked for the database, it said it wanted the state tape. It was given the county's tape instead. The tape was a print image and needed to be parsed. The paper asked again for the state tape and was denied. The state said that fortune hunters would read the papers, tell people that the state owed them money and take 25 percent from the recipients just for telling them. Eventually, the state gave in and the newspaper received the tape, Stith said.

"What we found in our story was that a number of people were owed huge amounts of money and that the state had failed to make even a passing attempt to find these people," Stith said. "One elderly woman had $40,000 being held by the state. She didn't want it back. She felt the state had forced her to sell her land, and she didn't want any part of taking their money."

These are some of the ways that newspaper stories can be improved by using computer-assisted reporting.

The *News and Observer* has used computer-assisted reporting to cover more in-depth stories it normally runs and to answer questions that sources won't.

For example, the paper did a feature story that told people how fast they could drive over the speed limit without getting a ticket from state troopers.

"The troopers we talked to said they looked for a clear and substantial violation depending on traffic, weather conditions and other things. In other words, they'd never tell you," Stith said. "When we looked at 300,000 violations on the database it was clear that 99 percent of the violations were for driving over 10 mph over the speed limit. It just jumped out at you when we graphed the information."

When the paper looked at how much lobbyists spend to wine and dine legislators, reporters found that using databases saved them precious time. Instead of spending days hand-copying the records at the Secretary of State's office, reporters got the records on floppy disks and within an hour started analyzing them with Report Writer for the next day's story, Stith said.

"This was much cheaper and better than the old way," Stith said. "It has to be, or we won't do it."

The News and Observer won't do a story just because it involves the computer, because there's too much out there that needs to be covered, Stith said.

"Computer-assisted reporting is like a tape recorder—it's another tool for accuracy," Stith said. "But if there's a paper report that will answer your questions better than a computer database, then use it and leave the computer alone."