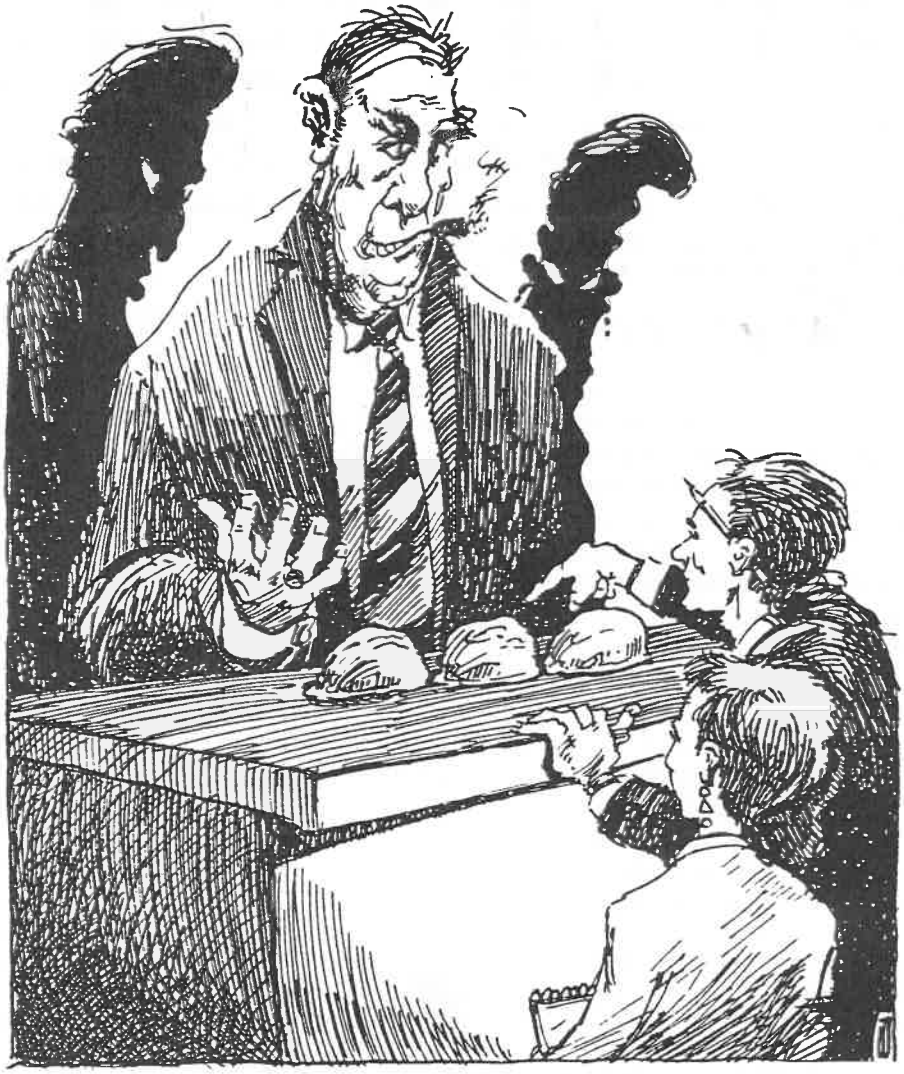
Beating them at their own Game



No longer able to baffle the press with talk of bytes and bits, bureaucrats are now using access laws to play the old shell game with reporters requesting electronic records.

Inside

The forum for computer-assisted reporting

August 1992 Vol.3, no. 8

How far will they go?

When Philadelphia Inquirer systems editor Al Hasbrouck wanted four years of criminal court dockets from the city, he had to fight for them. Then he had to fight for a copy of the codes for the more than 17,000 lawyers in the Philadelphia area, which he got in the form of a six-inch thick printout. Then he had to continue fighting for the codes on ninetrack magnetic computer tape, the way he had originally requested the records.

Boston Herald staff writer Dave Armstrong was familiar with the old-fashioned method of record redaction: the black pen marks that eliminated the parts of a paper record a journalist was not allowed to see. But Armstrong was not prepared for the new redaction "logic" that would prevent him from seeing any part of a computer tape, even the fields that hadn't been redacted.

By Lou Grieco
MICAR

Journalists are not alone in recognizing the potential of computer-assisted reporting. Government bureaucrats with whom reporters must deal to get tapes also see what journalists can do with personal computers and are learning new ways to deny or delay access to electronic records

Sometimes the resistance comes in an unsophisticated bluff, in which the reporter is told the records are not public, in the hope he or she will simply go away. But new strategies, some depending on various state access laws, can keep reporters busy just trying to get tapes.

"They've gotten real creative," Hasbrouck says. "Public servants understand what we can do to embarrass them through computer records."

Pennsylvania law is unclear on how public records are to be provided, and Hasbrouck says bureaucrats have stalled him in several ways. One time, an agency gave him a record layout in COBOL code, which Hasbrouck is unfamiliar with, so he needed to get a translator.

In the case of the court dockets, Hasbrouck had to go through endless negotiations. After he "finally wrung it from the hands of the good civil servants," he spent several weeks to get the codes he needed to understand the dockets.

Even in states where the requestee cannot determine what form public records are released in, reporters can still get the runaround. In Massachusetts, where the law says bureaucrats must give records in th forms requested if possible, interpretation makes the difference.

"There's one trick in particular we have a problem with," Armstrong says describing the redaction ploy. Another part of the Massachusetts law says that bureaucrats do not have to create records for reporters, just turn over what they already have. Certain Massachusetts agencies have decided that redaction in computer records creates new records, because of the programing required to provide records with certain fields deleted. So, the agencies say, the reporter cannot see the original records, because of the confidential fields. However, the reporter also has no right to see the records without the confidential fields because the "new" records aren't covered under state law.

"It's interesting, and at the same time it's bothersome," Armstrong says. The Herald does have a court case pending that will address this interpretation. Armstrong says he's optimistic about winning, though it may be the start of a long legal battle.

Jerry Uhrhammer, a reporter with the Tacoma (Wash.) Morning News Tribune, is skeptical of a new state public access law. Under the law, which went into effect earlier this year, Washington agencies have five days to respond to requests. The previous law mandated a non-specific "promptly."

Before, when he could insist on a prompt response, Uhrhammer says he often got answers within 24

MICAR
Is interested in any information, ideas or stories related to computer-assisted reporting for future issues of UPLINK
If you wish to contribute, please mail stories or ideas to:
120 Neff Hall
University of Missouri
Columbia, Mo. 65211
(314) 882-0684

hours of his requests. Now, with the new law in effect, Uhrhammer says the five-day rule may actually slow his requests down. He is finding that bureaucrats often wait until the last possible day to start working on a request.

"[The] most common stall is the delay, the foot dragging," Uhrhammer says. For him, the new law may give more excuses for foot-dragging officials.

Texas access laws were good, until electronic records became predominant. The state can determine the form the records will be provided in. Allen Pusey, assistant projects editor for the Dallas Morning News, says that many times, the paper receives huge printouts of records.

"I think that's been done specifically to make it harder," Pusey says. The Morning News uses data entry to create a database if it receives paper records. "Until we fix the law, it's the only thing we can do."

Cost is another factor that agencies can use to keep reporters at bay. In Bucks County, Pa., Hasbrouck found that agencies charged six cents per record, which is good for lawyers who want specific records but can get costly for reporters interested in spotting broad trends. Hasbrouck notes that it doesn't cost any more to store six or 6,000 records.

"A lot of times it's like taking a sledgehammer and banging them over the head."
-Dave Armstrong

In Washington, Uhrhammer has also found cost to be a problem. One agency offered to provide tapes, but wanted to charge $40 per hour of programming, and said it the duplication will take 40 hours. Uhrhammer says this is "outrageous and it's not legal." Under Washington law, agencies can only charge the cost of reproduction, and Uhrhammer says both the price and time estimates are way too high. The matter has not been resolved yet.

But regardless of new strategies, the favorite bureaucratic tactic remains the bluff. These battles are often easily won in court, but can be costly and time-consuming. Letters are exchanged by various attorneys, and if the requestor consulted the access laws before making the request, the bureaucrats will give in.

"A lot of times it's like taking a sledgehammer and banging them over the head," says Armstrong.

Pusey has his own strategy for waking up the civil servants. The Morning News has a well-known policy of pursuing legal fees as well as access. "That I think gets their attention." While Pusey concedes that some bureaucrats are getting more sophisticated about computer-assisted reporting, many will rely on the bluff because they don't know the access laws and don't care. ■

A Few Tips on Making Computer Requests

By Sandra Davidson Scott

1. Make requests as specific as possible and direct them to the designated custodian of that particular type of record. Remember that the custodian may need computer experts to handle your request, so requesting a meeting with both the custodian and his or her computer experts at the outset may speed up the process. At the meeting you can discuss such issues as how much of a computer expert's time it would take to retrieve your request.

2. Take time to know your state's computer laws relating to your request before the meeting, including:
* Exemption laws. These laws which exempt certain records from public disclosure are sometimes scattered throughout the statutes instead of being contained within the chapter or title on access to public records. Be prepared to argue why your request is not exempt.
* Redaction laws. If a record you request contains some information that is exempt and some that is open to disclosure, does your law say that the custodian must delete the exempt information and give you the rest? If not, you may have to argue for redaction. (Note that the United States Supreme Court did not rule in favor of redaction of F.B.I. records in Department of Justice v. Reporters Committe for Freedom of the Press, 489 U.S. 749 (1989).)
* Tailoring laws. Do your statues say that custodians must provide records in computer form or, better yet, in the format you request if the computer system is capable of doing so? If your statues don't address tailoring, you may have to argue that what you are requesting does not constitute creating a new record. "We don't have to do programming" may be an argument you'll face. Think in terms of the ease with which your request can be fulfilled: you may have to argue that a few computer strokes does not constitute programming.
* Costs. If your statute reads in terms of "reasonable" or "actual" cost, be prepared to argue precisely what that imprecise language means. You want to limit the costs to costs of search and copying at a maximum, and you want to limit the costs to costs of search and copying at a maximum, and you want to avoid paying for overhead.
3. Think twice before filing a lawsuit. If you have "fuzzy" statutes, you may be better off living with uncertainty and arguing with custodians than living with a bad court decision. For instance, it may be better to have no court decision if you have statutes that do not address the question of redaction than to have a court decision that says custodians of records do not have to redact records ■

Connections that pay off

St. Louis Post reporters use computers to uncover scandal involving state worker's comp laywers and Missouri Attorney General.

By David Raziq
MICAR

For years, few Missourians had heard of the state's Worker's Compensation Second Injury Fund. Created by a tax on insurance premiums, it gave workers the opportunity to sue the fund for extra disability when an on-the-job injury aggravated a previous injury or medical condition. The fund was created to ensure that workers got a square deal from worker's comp. However, by early 1992 some Missouri lawyers had voiced complaints about they way in which the fund was being administered.

"A couple of attorneys who had worked on the fund had been calling us for months," says reporter George Landau of the St. Louis Post-Dispatch,"They were disgusted with the way it worked."

And how it worked, claimed these attorneys, bordered on the illegal. Fifteen private lawyers had been hired by state Attorney General William Webster to defend the second Injury fund. However, Webster also allowed these lawyers to bring suit against the same fund on behalf of their own clients. Hence, these "Special Attorneys General" would negotiate settlements among themselves, one lawyer suing the fund one day while another lawyer would defend it. Later, they would switch roles.

In addition, the Post's sources claimed that these lawyers or their firms had made contributions to Webster's campaign for Governor. Hence, claimant's lawyers who had also contributed allegedly received better settlements for their clients.

After receiving the complaints, the Post launched an investigation and planned to use a computer to create a database that combined election contribution data with records from the second injury fund to support the allegations.

But, like many states, Missouri's state election contribution data is not computerized, but kept in a hardcopy format. State law also requires agencies to charge 50 cents per photocopy which can become prohibitively expensive.

However, foresight and a little serendipity nipped these problems in the bud. "Post reporter Terry Ganey and the capital bureau had campaigned for a few years for a database of campaign contributions," Landau said. So when the story came along, an office manager had already been typing the contribution reports into a Foxpro data entry program. Plus, the Post bought the reports in a microfilm format for five dollars per reel, with each reel containing two thousand pages.

"We were just lucky," says Landau, "because she had just finished the Webster campaign contributions from 1988 to 1991 when I got the second injury fund database from the Division of Worker's Comp."

Acquiring the database, which listed the names of attorneys filing claims, attorneys defending the fund and approving settlements, and the settlement amounts, was a difficult task. The Post team had to negotiate with the Attorney General's office for the tapes. At first the Attorney General's office gave the Post the runaround saying they didn't have the information the newspaper had requested.

Ganey, Landau, and the Post lawyers decided that it was time to give the bureaucrats an attitude adjustment. "This involved a lot of hasseling and discussions between our attorney and the Attorney General's office." And five weeks later: "They finally agreed we were entitled to the electronic data."

On the nine-track tape Landau received was a "print-image" of a print-out the Post had previously aquired, with column headings and page numbers. "The data had the same position in each line of the file," says Landau, so he just edited out those lines containing the page-numbers and headings. The result: a fixed position ASCII document that he could now convert to Foxpro. It was time to get to work.

"The hard part of a project like this is matching names because they are spelled differently in each database," says Landau," so I developed some little tricks." For example, in the second Injury database had dozens of attorneys with more than 20 cases, each with their names spelled differently from case to case. "So, what I would do," says Landau, "is sort the names alphabetically into a new table and the like-names would usually cluster.

The Post's Mission: Prove that private lawyers hired by Missouri Attorney General William Webster to defend the state's second injury fund and lawyers who made contibutions to Webster's campaign received better settlements than attorneys who did niether. Using computer tapes from the State Division of Worker's Compensation and a special state contribution database, the Post discovered:

***Attorneys who didn't contribute to Webster nor defended the fund received an average award of:** $3,548

***Attorneys who did contribute to Webster's campaign received an average award of:** $5,307

***Attorneys hired by Webster received an average award of:** $11,808

Once I was sure this cluster was the same person, I put a 'plus' in front of the first name-occurence and a 'minus' in front of the last name-occurence. Then I would write a program that replaces a new field in the original database with a consistent name, usually the first name in the cluster. So now everything was spelled consistently."

Landau says he applied this technique throughout the data, saving him "hours and hours of time." Then he wrote another Foxpro program that allowed him to link the record numbers in the Webster contributers database with the record numbers of the now consistently-named attorneys in the second Injury database. Another program then went through both databases, looked at each record number, and aggregated all the contribution data for each lawyer into the 2nd Injury table. "That may be a complicated way to do it" says Landau,"but I wanted to end up with a single table that told the whole story."

However, Landau says it took more than a computer table to tell the story and praised his partner's reporting skills. "I did all the computer stuff but Terry Ganey did some really dynamite reporting by working her sources. And if it hadn't been for those sources saying how the fund worked, we wouldn't have had as clear an idea of what we were looking for."

Landau continues: "It can be hard to analyze a database and if you don't know what you're looking for the story won't jump out at you. Thats obvious, but people can forget it when you do computer stories.■

Tech Tips Tech Tips Tech Tips Tech Tips

By Elliot Jaspin

Problem: A database that lists people and companies is nearly impossible to work with because of variations in spelling. A company such as Amalgamated Zoot Suits might be listed as:

Amalgamated Zoot Suits
Amalgamated ZootSuits
Amalgamated Zoot-Suits
Amalgamated Zoot Suits Inc.
Amalgamated Zoot-Suits Inc.
Amalgamatd Zoot Soots

Because the database is large, Correcting the spelling of every entry is not an option.

Solution: If we look at the variations in the spelling of Amalgamated Zoot Suits, notice that changes occur somewhere in the last half of the name. We can make use of that fact to create a new name field that uses only the first half of the name field. This is done in XDB by using two commands: Update and Xleft.

The first step, of course is to create a new field in our hypothetical database called Badspell that will hold the new name we are creating. Let us assume we are going to use the first 11 letters of the name field. Thus the "NewName" field will be a "char" field 11 characters wide.

We then run the following command:

```
Update Badspell
set NewName = xleft(name, 11)
```

We can now use the "NewName" field to group, sort and order our database. Or can we?

While the "NewName" field is an improvement, it is not a perfect solution. Notice that the last entry for Amalgamated Zoot suits contains a misspelling with in the first 11 letters. If we try to group on the "NewName" field, "Amalgamatd" will be seen as a different company. There are two possible solutions: grab fewer letters to get around the misspelling or correct the origional entry and rerun the update command.

Reducing the size of "NewName" might work except that there is also a person in the database named "Amalgam, Arnold." If we take too few letters, "Amalgam, Arnold," and "Amalgamated Zoot Suits" might wind up being lumped together.

It might be simpler to run the following command:

```
Update Badspell
set NewName = "Amalgamated"
where xleft(name,11) ="Amalgamated"
```

However, there also might be two companies who start with the word "Amalgamated." To account for that we could make "NewName" field larger and grab more letters from name or, alternatively, keep NewName the same size and use the Update command to change "Amalgamated" to "Amalgam Zoo" wherever the first 13 letters in "name" field are "Amalgamated Z." The second approach would leave us with two different "NewName" entries:

"Amalgamated" for "Amalgamated Federated Industries"
Amalgam Zoo" for "Amalgamated Zoot Suits"

While we may face several more runs to account for such problems, this approach accomplishes two things: It "Solves" the large majority of misspellings on the first pass of the database and it allows us to "Correct" any other misspellings in subsequent without actually having to go into the database and make the individual changes.■

What they didn't teach you in school

By William Syken
MICAR

So you want to start a computer-assisted reporting program. Maybe you've read some books, or taken a seminar on how to use the equipment. Now you're ready to give it a whirl in the newsroom.

No problem, right?

Well...

Most reporters said there was a huge difference between the real world and the controlled atmosphere of the seminar classroom.

Jerry Uhrhammer of the Morning News Tribune in Tacoma, Washington said the learning curve was steeper than what he expected after completing a training seminar at the Missouri Institute for Computer-Assisted Reporting in Columbia, Mo..

"Learning how to do this is like learning a foreign language, and you don't learn it in two weeks," Uhrhammer said. "You need to immerse yourself and speak it constantly. I'm still in the learning process."

Uhrhammer also said the Morning News Tribune's computer-assisted reporting program got off to a slow start because, as a money saving maneuver, they tried to get their program running with a Vax business computer.

"When we found that was not working we went out and bought the hard drive and the 486," he said.

At the Akron Beacon Journal, Bob Paynter's also found that his paper's initial equipment purchase designed to save money soon proved inadequate. The Journal started out with the Qualstar (1600 BPI). Now he has the Qualstar 6250, but still finds the going slow. Paynter said his biggest problem has been setting up a user-friendly way to get reporters into databases. Paynter doesn't have programming background, and has struggled to make his datasets accessible.

"I've aquired a lot of databases that would be valuable to reporters on a daily basis," he said. "How do you spread it around without going through a Jaspin-like seminar?"

Some papers complain they simply haven't had the opportunity to use their equipment.

Rae Davidson of the Blade in Toledo, Ohio, said tight staffing at her paper had prevented people from committing themselves to in-depth computer-assisted projects.

"The reporters we have are very interested," she said. "But this is the nineties. We may be in the age of technology, but we're also in the age of doing more work with fewer reporters."

Uhrhammer also said he hadn't used the equipment as much as he would like, but cited a different reason. He said after getting the drive he has been involved with investigations that either didn't involve computers or had been based from information on diskette rather than 9-track tape. He has only used the tape drive to download census tapes.

Judy Nichols of the Arizona Republic who took the MICAR seminar in January, said she has had an almost trouble-free experience with her computer-assisted program and credits Russ Beutner, a former MICAR employee and Republic intern for her success.

"But even with Russ here, it has been a learning experience for me and the editors," Nichols said. "When you start to do analysis, it can get more in depth than you expected rather quickly."■