May 1998

Uplink

A newsletter of the National Institute for Computer-Assisted Reporting

HEALTH CARE INVESTIGATION

Wrestling giants

By Josh Barbanel
The New York Times

We were gathered in a tiny windowless conference room at *The New York Times* noodling over what was undoubtedly going to be an epic tale of the troubled American health care system. But which epic would it be?

For weeks, Martin Gottlieb, Kurt Eichenwald and Tamar Lewin had traveled around the country meeting with sources, reviewing court and agency records, exploring the troubling business strategies that turned Columbia/HCA Hospital Corp. into a $20 billion health care giant, the world's largest health care provider.

After a few days on the road in places like Tallahassee and Baltimore, I stayed back at the office staring at my computer screen. I was trying to make sense of a carton filled with 89 computer cartridges containing more than 30 million medical billing records perched on the edge of my desk.

As we reviewed our findings – tales of hard-ball takeover practices, questionable incentives for doctors to switch hospitals and bring their patients with them – we struggled over fundamental questions of fairness.

Was this a story about a greedy for-profit company that went way too far? Or a story about the sorry state of the entire American health care system? Was Columbia/HCA a special case or were not-for-profit hospitals, facing the same pressures, doing exactly the same things? Were we picking on the wrong guy?

Analytical evidence

At this pivotal moment in the project, I was able to lean forward on the sagging couch and pull out the computer-assisted reporting analysis of medical billing records. I argued that they provided the strongest and clearest evidence yet that Columbia/HCA was simply different from other hospitals and needed to be treated that way.

The analysis showed that Medicare recipients who were treated at Columbia/HCA hospitals in Texas cost the government millions of dollars more than patients treated at other hospitals. Their costs were higher than costs at all other hospitals, particularly when the entire cost of care included the hospital stay through skilled nursing, rehabilitation and home health care. Their costs were far higher than the costs at comparable not-for-profit hospitals.

Continued on page six

INTERSTATE TRAPPINGS

Sudden Seizures

By David Washburn
The Morning Call
(formerly Dateline NBC)

Dateline NBC's IRE award-winning investigation of a Louisiana sheriff department's illegal use of drug forfeiture money – "Probable Cause" – started with a ski trip in Aspen. It ended with a cardboard box filled with photocopied checks and a Microsoft Access database.

It was a week-long luxury ski trip. Four Jefferson Davis Parish police deputies rented a condo, bought ski clothing in ritzy boutiques and took to the slopes. Police billed the trip as "deputy training."

All expenses were paid with money police seized from alleged drug dealers in Louisiana, mostly along Interstate 10.

Louisiana Legislative Auditor Daniel Kyle investigated this 1990 ski trip and found that not only was it bogus but an illegal use of drug forfeiture money. An additional $200,000 was missing. One

Continued on page eight

WINNING WAYS

Update

The envelopes have been opened and the trophies shelved. Now it's time for the award winners to share their CAR techniques.

In this issue, Josh Barbanel of *The New York Times* revisits his paper's computer-assisted, statistics-laden analysis of health care behemoth Columbia/HCA and its aggressive, profit-seeking schemes. David Washburn reviews Dateline NBC's uncovering of asset forfeiture abuse by law officers in one Louisiana parish. Stuart Watson of WRAL-TV reveals the tip that led him to investigate the intrinsically flawed military health care system that protects malpractice. Dan Keating of *The Miami Herald* details his paper's foray into the excessive practice of police officers loading up on hours of court appearances for overtime pay.

And, in the ongoing discussion of the value of intranets for newsrooms, Ray Robinson of *The Virginian-Pilot* and Tom Boyer of *The Seattle Times* check in with their take on the software Cold Fusion.

Uplink

May 1998
Volume 10, Number 4
A newsletter of the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston
brant@nicar.org

MANAGING EDITOR
Brent Johnson
bjohnson@nicar.org

SENIOR CONTRIBUTING EDITORS
Len Bruzzese
len@ire.org
Sarah Cohen
sarah@nicar.org
Richard Mullins
richard@nicar.org

COPY EDITOR
Jeanine M. Davis

ART DIRECTOR
Wendy Charron

STAFF
Jack Dolan
Seth Hemmelgarn
Justin Mayo
Debbie Roberts
John Sullivan
Alistair White
Ted Peterson

Uplink is published every month by the National Institute for Computer-Assisted Reporting, 138 Neff Hall Annex Columbia, MO 65211. (573) 882-0684. Subscriptions are $40 for IRE members, $60 for nonmembers. Postmaster: Please send address changes to NICAR. Send e-mail to jeanine@nicar.org NICAR is a joint effort of Investigative Reporters and Editors and the University of Missouri School of Journalism. NICAR services include hands-on newsroom training on computer-assisted reporting, special academic and advanced training in data analysis. NICAR is supported by grants from The Freedom Forum and other foundations intended to help the institute deliver its services nationwide to news organizations and associations.

System protecting military malpractice

By Stuart Watson
WRAL-TV

Stuart Watson attended the NICAR basic boot camp in August 1995.

Normally I'm terrible at phone tips. To my detriment. If Deep Throat called, I'd put 'em on terminal hold, then cut 'em off. Maybe that's why my colleagues rib me about being more at home with a stack of documents or a database. But it was hard to hang up on Kristie Piittmann. She called March 22, 1995, to say Army medicine killed her baby.

Having firsthand witnessed labor and delivery four times, her story hit home. You could see it happening. The fever. The racing fetal heartbeat. The antibiotics administered too late. The awkwardness of telling the mother when Cole Piittmann was born a few weeks before Christmas 1994. Then the rush to a civilian hospital and the parents' ultimate choice: handing over the infant to be airlifted to Duke University with little hope of survival or unhooking all the tubes and letting their only son die in their arms. They chose the latter.

It was a heart-rending story, all right, the news director said, but how was it different from allegations of malpractice in a civilian hospital among civilian doctors? In other words, you have a story – now what is the report? What is the meaning, the context? The full answer would take two and a half years of reporting.

Alley cats

We went down a number of blind alleys before finding the lessons of Cole Piittmann's death. First we checked the death rates at the military hospital against death rates in the neighboring civilian hospital. Fewer infants died at the military hospital. Then we checked the status of infants transferred from the military hospital to civilian hospitals using a database of death records downloaded from the Web. Finally we checked the number of infants at various hospitals who died of Group B strep infections like Cole Piittmann had. The alleys led nowhere.

The next set of blind alleys concerned the doctors. We filed Freedom of Information Act requests asking whether Kristie Piittmann's doctors had ever been sued before. The Army denied us saying it couldn't find such information, that doctors' names aren't linked to malpractice claims, and even if they were, the information was exempt from public inspection.

We also asked whether the doctors had ever been disciplined. Again denied. An Army doctor's disciplinary history – short of court-martial – is exempt from public disclosure. But the U.S. Army Claims Service at Ft. George Meade in Maryland mentioned a database. Did somebody say 'database'? We filed an FOI request.

The Army withheld names and identifying information for claimants. But after one appeal, the Army included the names, addresses and phone numbers of attorneys. Some attorneys proved helpful in tracking down old clients. The U.S. Army CLAIMS database arrived in the form of a tiny cassette not much bigger than an audio cassette. We didn't have a player. After consulting CAR gurus on the NICAR listserv, we hired a local data house to transfer the cassette to nine-track tape. Though we didn't have a nine-track tape drive in house, our corporate data processing officer did. We begged him to clean it up and FTP it to us.

All of this took months. The corporate data processors would ask little questions like, "Do you want us to clean this Army data or do you want your paycheck this week?" In the end, we had a complete database of every tort claim filed against the Army for about eight years.

This database could have told us the type and rates of malpractice claims at various military hospitals. But there were all sorts of caveats. Some hospitals were much larger. Some might have been in areas where more aggressive trial attorneys recruited and encouraged military families to sue. In the end, we primarily used the data to track down similar malpractice cases at the same facility.

Continued on page three

Continued from page two:

Disordered care

Civilian differences

Then I bumped into Russell Carollo, an investigative reporter at the *Dayton Daily News*. I had called Carollo in the summer of 1996 to ask how he got information about missing and stolen military weapons. He had shared his data, provided that we credited his news organization, which we did.

The military protects bad medicine. Military doctors do not carry their own malpractice insurance like civilian doctors, are not sued like civilian doctors and are not disciplined like civilian doctors.

I mentioned that I was interested in medical malpractice in the military. He said, "Stuart, the stars have aligned. We've been working on this for months." Carollo and Jim Stewart at CBS News helped sharpen our focus. They said they didn't think we would ever be able to fairly compare malpractice rates at military and civilian facilities. They said the systems were too different.

The story, they said, was the differences between the two systems and how those systemic differences left military families unprotected against malpractice. Foremost, active duty soldiers, sailors, airmen and marines could not even file malpractice claims. Doctors were not named individually in claims. Military doctors didn't carry their own malpractice insurance. And disciplinary records were hidden. (We knew this last fact from our FOIA experience.)

Carollo and Jim Nesmith at the Cox News Service had already done considerable work. Carollo offered to share if WRAL didn't air the report before the *Dayton Daily News* went to print.

Carollo had accumulated databases of military doctors from all branches. We later split the cost of hiring a data entry house recommended by a NICAR listserv member to input the Public Citizen list of more than 13,000 "questionable" doctors and run it against the databases of military doctors.

We then shared the grunt work of filing open records requests with various state medical boards to ensure the doctors mentioned in the Public Citizen list were the same ones in the military databases.

Research reaction

The years of research paid off in October 1997 when WRAL-TV, the *Dayton Daily News* and the Cox News Service reported how the military protects bad medicine.

Military doctors do not carry their own malpractice insurance like civilian doctors, are not sued like civilian doctors and are not disciplined like civilian doctors.

The Pentagon does not report to the National Practitioners Data Bank the same information as the civilian world does.

And, unlike civilian doctors, military doctors can continue to practice after repeatedly failing state medical exams.

The Army Times, Navy Times, Air Force Times and wire services all ran Carollo and Nesmith's series. WRAL aired a five-part series and a half-hour special report.

The Assistant Secretary of Defense for Health Affairs responded with a 13-point plan for improving military health care. The Surgeons General of the Army, Navy and Air Force were called to a closed-door meeting with the National Security Subcommittee of the House Appropriations Committee. The subcommittee later set aside $5 million for an expert panel to evaluate the quality of military health care.

The Pentagon pulled doctors who failed medical exams from clinical positions.

And it all started with a phone call, a massive database, and lots of teamwork with NICARians!

Stuart Watson can be reached at
(919) 821-8705 or by e-mail at
stuart@interpath.com

WRAL's investigation can be accessed online at www.wraltv.com/news/wral/5investigates/1997/1005-womack-part1

In addition to an IRE Certificate, WRAL-TV won the coveted George Foster Peabody Award – the only local commercial station in the country to do so.

Carollo and Nesmith of the *Dayton Daily News* won the George Polk Award and the Pulitzer Prize for National Reporting for their series, which was reviewed in the Dec. 1997/Jan. 1998 Uplink.

Undeserved overtime

By Dan Keating
The Miami Herald

In late April, Metro Dade Police – the largest agency in south Florida – punished 36 officers, the largest disciplinary effort in its history. Supervisors were demoted and officers were suspended without pay for a widespread scam in which officers, even if they had not performed any police work, added their names to arrest reports in order to collect court overtime.

It's usually hard to turn multiple records into individual summaries with such a wide variety of information. We used SAS because it can manipulate a group of records to produce a single summary record.

The department's nine-month internal investigation was prompted by *The Miami Herald's* 1997 investigation, "Collars for Dollars." In April, the project was named a Pulitzer finalist in the investigative category.

The story pulsed with rich anecdotes produced by the shoe-leather effort of three reporters on *The Herald's* investigative team. Its foundation lay in ironclad assertions drawn from data analysis, which involved nine different criminal justice datasets from a variety of sources.

The key ingredient was the court's electronic docket that lists officers subpoenaed to court. We obtained records of 704,077 officers subpoenaed in 143,055 felony, misdemeanor, juvenile and traffic cases in Dade County from 1994 through 1996. Our main goal was to tally how many officers were coming to court and then calculate by case, agency, officer and charge.

The hard work wasn't analyzing the data – it was organizing it.

Unfortunately, the raw data was structured in a format that couldn't be used for effective analysis. The database arrived in two tables. One table listed the case number, charge and disposition. With separate records for every charge, each case could have a dozen records or more. The second table had one record for each subpoena, listing the case number, agency, officer and hearing date. It, too, could have a dozen records per case.

Although the tables were related by case number, any attempt to join them produced a mishmash of up to 100 records for an individual case.

The multiplicity challenge

Before we could do anything, we had the tricky chore of reshaping the data. We needed a single record to summarize each case, but we couldn't lose the ability to pinpoint specific officers or charges.

For the charges, we made a single summary record for each case. The record included the first eight charges (if that many were included), a count of the total number of charges, and dispositions.

To be able to quickly find particular kinds of cases, we created yes/no flags for several types of charges: from murder and robbery to prostitution and drunk driving. If a case involved any of those charges, its flag was set to "yes." Since the project concentrated on misdemeanors, the flags let us quickly run statistics on those cases.

We also created a single record for each case in the subpoena database. It included the total number of officers subpoenaed, the names of up to three departments and how many officers were subpoenaed from each. It also noted the officer and department that initiated the case. That showed us how groups of officers piled onto cases.

It's usually hard to turn multiple records into individual summaries with such a wide variety of information. We used SAS because it can manipulate a group of records to produce a single summary record.

To be able to find specific officers, we kept a dataset listing one subpoena per record. A name search would draw up all the case numbers for a given officer. The case numbers could then be linked to the charge sum-

Continued on page five

THE "COLLARS FOR DOLLARS" SERIES, WHICH WAS A PULITZER FINALIST, CAN BE ACCESSED ONLINE AT WWW.HERALD.COM/ARCHIVE/COLLAR

MOST OF THE ANALYSIS FOR THE STORY WAS DONE ON A PENTIUM 166 WITH 64 MEGS OF RAM RUNNING OS/2 WARP 4.0, SAS FOR OS/2, ACCESS 2.0 AND EXCEL 5.0. THE DATA FILES OCCUPIED ABOUT 325 MEGABYTES.

Continued from page four:

Cops a-courting

mary or subpoena summary to pull statistics for that officer.

Once we had the SAS program to reshape the felony/misdemeanor data, we could change file names to run the same operation on the juvenile data and then the traffic data. That ability to recycle SAS code and repeat a lengthy series of steps came in handy since we ended up getting four sets of the traffic data before the court folks finally produced a database including all police witnesses on a case.

In any analysis, a baseline for comparison is the most important part: What can we use to define what's normal or abnormal?

Gone fishin'

With the data reshaped, it was time to actually analyze it. First, we cast a sweeping net through the data to catch tips (cases and officers) for the reporters to track down.

We specified a kind of case using the yes/no flags and then grabbed the cases with the most witnesses. We also pulled the officers whose cases averaged the highest number of witnesses.

To see who was in the habit of joining with a friend, we checked how many times any two officers appeared together on a case. To do that, we made a temporary duplicate of the subpoena dataset in order to compare it with itself.

With help from my predecessor and mentor, Professor Steve Doig of the Cronkite School of Journalism at Arizona State University, we used the following code. It's SQL code, but for this story run under SAS. You might find it useful anytime you're trying to match a dataset against itself to pull out pairs.

The original dataset was named subp1 and the duplicate subp2. In the duplicate, the officer identification was renamed copid2 and the case number was renamed casenum2. The code uses a "less than" on the officer identification number so that each pair can be listed only one way (you won't have both "Keating-Doig" and "Doig-Keating"):

```
create table witness.subp3 as
 select subp1.copid,subp1.casenum,
       subp2.copid2,subp2.casenum2
 from subp1, subp2
 where subp1.casenum =
       subp2.casenum2
 and copid lt copid2
 order by 1, 2;
```

The code produced a table of matched pairs with case numbers. We then grouped by the pairs and counted to find which officers appeared together the most. We could also see all the cases of any given pair.

Using those techniques, we found plenty of cases worth checking out. At least six police witnesses appeared in 128 misdemeanor prostitution cases. Two officers (close friends) were listed 238 times as witnesses in each other's drunken-driving cases, even though they worked for different departments in different jurisdictions.

Publication precision

To produce statistics for publication, our analysis had to be more precise. We considered other complicating factors, such as officers subpoenaed only to testify about the calibration of the breathalyzer, the tiny proportion of cases with guilty pleas at arraignment, or cases with multiple defendants.

In any analysis, a baseline for comparison is the most important part: What can we use to define what's normal or abnormal? In our case, we had four major police departments, three that pay for court overtime and one that does not.

That made it easy to draw comparisons. For instance, the department that doesn't pay for overtime was 100 times less likely to have at least five officers on a DUI case. The three departments that paid overtime had 6,445 routine DUI cases with at least five officers in three years. Experts said three officers would be the maximum needed to prosecute a case.

One piece of telling evidence: the convic-

Continued on page thirteen

UPCOMING BOOT CAMPS:

JULY 12-17, 1998 – BASIC BOOT CAMP IN COLUMBIA, MO.

AUGUST 9-14, 1998 – BASIC BOOT CAMP IN COLUMBIA, MO.

AUGUST 14-16, 1998 – ADVANCED BOOT CAMP IN COLUMBIA, MO.

FOR A COMPLETE LISTING OF UPCOMING IRE AND NICAR EVENTS, INCLUDING BOOT CAMPS, ON-THE-ROAD TRAINING, SEMINARS, AND REGIONAL CONFERENCES, VISIT OUR WEB SITE AT WWW.IRE.ORG/CALENDAR.HTML

THE "HEALTH CARE'S GIANT" SERIES, IN ADDITION TO BEING AN IRE FINALIST, WON THE GEORGE POLK AWARD FOR BUSINESS REPORTING AND WAS A FINALIST FOR HARVARD UNIVERSITY'S GOLDSMITH AWARD.

TO STEER YOUR CAR PROJECTS IN THE RIGHT DIRECTION, ORDER "COMPUTER-ASSISTED REPORTING: A PRACTICAL GUIDE" BY BRANT HOUSTON. IT CAN BE ORDERED FROM NICAR AND INVESTIGATIVE REPORTERS AND EDITORS FOR $26 FOR IRE MEMBERS OR $30 FOR NON-MEMBERS PLUS SHIPPING. CALL (573) 882-2042.

Continued from page one:

Bold profiteers

And why not? Columbia/HCA ran both the hospitals and the home health care agencies – and everything in between. They were aggressively trying to maximize their profits.

We didn't know it then, but we were still at the early stages of an investigation that would take more than a year before the first word made it into print. And the computer-assisted analysis kept the project focused squarely on Columbia/HCA and its practices. Over the ensuing months, as more allegations and information surfaced, we were able to use the same approach again and again. The interplay of data analysis and investigative hunches and leads brought this project to life.

The interplay of data analysis and investigative hunches and leads brought this project to life. The data analysis allowed the team to move from a specific allegation to an examination of an entire hospital conglomerate and how the government regulated it.

The data analysis allowed the team to move from a specific allegation to an examination of an entire hospital conglomerate and how the government regulated it.

Share-raising story

Eventually we pointed to problems and possible illegalities in many important areas of Columbia's operation that stemmed from the company's single-minded quest for profits and an ever higher price per share.

We looked at schemes to overbill Medicaid for millions of dollars per year, evidence that Columbia executives promised jobs and incentives to administrators of hospitals it was trying to buy, and data suggesting that some doctors who joined partnerships with Columbia skewed their admitting practices to enrich their investments. There were stories on the misuse of foundations set up when Columbia purchased not-for-profit hospitals and computer-assisted evidence that administrators inflated cost reports to the government to increase Medicare payments and cut care to the uninsured poor after a hospital purchase.

And finally, while Columbia/HCA had argued that its cost-cutting had led to lower prices for patients and thus a benefit to the public at large, the investigation found that care at Columbia hospitals actually cost significantly more than care at comparable institutions.

Writing about Columbia/HCA required us to connect dots that most often had remained unconnected. It was uncharted territory, at least for us, and sometimes information that we came across in the first few weeks didn't make sense or assume importance until months later.

The computer-assisted reporting grew out of investigative hunches we kicked around at the very start of the project. Columbia/HCA had consolidated whole segments of the health care industry and increased their profitability at the same time.

Troubling episode

We decided to look at the entire cost of care whether provided by Columbia/HCA-owned hospitals, skilled nursing units, home health care units or doctors who had invested in Columbia/HCA.

For Medicare patients, the Health Care Financing Administration keeps all the records and bills and gives researchers access to some of them for medical and policy research. We obtained separate billing records for physician visits, hospital stays, skilled nursing care, medical equipment and home health care. These records were all linked together by an encrypted patient Medicare number.

But what could we do with these records?

Continued on page seven

Continued from page six:

Billing inflation

My approach in this situation is to find researchers who have done parallel medical or social science work, then adapt and extend their methods to solve a reporting problem using administrative records.

This time, I found that there was a body of work in the medical literature examining "episodes" of care — what happens in the course of medical treatment both in and out of the hospital.

Our team used the Times mainframe and ran through millions of patient records. We combined all the bills and calculated the entire cost of an episode of hospital care from the day of admission until 30, 60 or 90 days later.

To make these comparisons fair and reasonable, we adjusted the cost figures for the differences in diagnosis and severity based on information coded in each hospital record. A variety of statistical methods were used to examine the statistically significant results.

As we suspected, Columbia/HCA hospitals had shorter length of stays during initial hospitalizations – where Medicare pays a fixed cost for a particular illness no matter how many days a patient is in the hospital – and longer stays in skilled nursing units – where Columbia/HCA is paid by the day. Costs were roughly the same or lower for Columbia for the initial stay. But they were significantly higher when the entire episode of care was included.

This work helped keep us focused on what was special and troubling about Columbia/HCA and its aggressive practices, rather than a more general hand-wringing approach about the changing health care system.

Applied training

It also trained us to work with the Medicare records and prepared us for more provocative findings later on.

Months later, we learned of allegations that a Columbia/HCA hospital in Tennessee was miscoding medical records to increase reimbursements from the Medicare system. There were four diagnosis codes for four different kinds of pneumonia, each with a different Medicare payment level. The Columbia hospital was far more likely to code to the highest paying – and most profitable – illness than surrounding hospitals.

Now that we were experienced hospital researchers, we applied that experience to the new allegation. Within a day or so, we showed that this pattern was repeated across Texas and Florida. When we ranked hospitals in both states on the percentage of cases coded to this high paying category, Columbia/HCA hospitals ranked at the top of the list in both states.

We had obtained four years of Florida state hospital records and were soon able to document that Columbia/HCA hospitals increased the proportion of cases coded to the most profitable category.

We also used annual financial and cost reports compiled by Florida and the federal Medicare agency. When we learned that Florida required hospitals to file reports showing all doctors who are investors and referred patients, we obtained them on paper and created a database. This allowed us to examine doctors who admitted patients to two or more hospitals to look for evidence that their investments influenced their practice patterns.

To process all that data, we used an IBM mainframe at the *Times* and were forced to learn how to do some complicated programming in SAS. Once we produced interim results, we brought them down to PCs and did more work in FoxPro, SPSS and various spreadsheets.

When Columbia suggested that the differences in costs had to do with hospital size or other factors, we ran regressions to show that the differences remained significant after accounting for hospital variables such as wage differences, size, and whether the hospital carried out teaching programs.

All this work produced an initial series of five stories and then 30 other stories by Kurt Eichenwald on government investigations. This in turn led to the ouster of Richard Scott, the chairman of Columbia/HCA and most of his top aides who had shaped the company's policies and personality. The new management team promised reforms that addressed, almost point for point, the major issues raised in the series.

Josh Barbanel can be reached at
(212) 556-7552 or by e-mail at
barbanel@nytimes.com

"Probable Cause" won an IRE Award.

All IRE Award winners are searchable through the IRE Web site, www.ire.org

NICAR is collecting stories that depended in part on statistical techniques – anything from simple crosstabs to averages versus medians to regressions. Sarah Cohen is particularly interested in stories in which the numbers never went into the paper or on the air but were used to help focus the reporting.

Send clips to Sarah Cohen at IRE, 138 Neff Annex, Columbia, MO 65211 or e-mail her at sarah@nicar.org with the appropriate link.

Continued from page one:

Forfeiture abuse

Louisiana police officer called the ski trip just the tip of an illegal iceberg.

Dateline producer Patrick Weiland had been keeping an eye on the story since 1994. He was hearing outrageous accounts of police officers pulling motorists over on I-10 and shaking them down for cash, seizing cars and property without any evidence, targeting minority drivers for traffic stops and misusing funds in the drug forfeiture account.

"We talked to a few people and the under-the-table stuff they were talking about was incredible," Weiland said.

Set up to shakedown

The key to the story was the law itself and how it differs from every other state. In most states, seized drug money goes into a statewide crime fund. In Louisiana, individual police departments get to keep 60 percent of everything they seize. And much of that goes directly to the officers who make the busts.

This gives local police ample incentive to seize as much money and property as possible. So much incentive, said Weiland's sources, that motorists traveling Interstate 10 were getting pulled over every day for bogus reasons by officers acting more like shakedown artists than cops.

Weiland filed a Freedom of Information Act request for five years' worth of checks written against the drug forfeiture account in Jefferson Davis Parish. He was granted access to the records, but only on paper. The parish didn't keep the records on computer.

In fact, when associate producer Lindsey Schwartz went down to the parish to get the records, she found 1,500 photocopies of checks stashed away in a cardboard box.

"The checks were coming back from the bank and just put in the box," Schwartz said. "No one ever looked at them."

After a few hours of going through the copies, Schwartz realized two things: She had the evidence they were looking for and she needed help.

We designed a simple table in Microsoft Access based on the check. It included the check number, the date of the check, the payee and the amount.

That was the easy part. The hard part was entering information from 1,500 checks.

Glazed assets

After three weeks of entering data, it took just a few queries to find out that police were spending seized money all over town.

Thousands of dollars were spent at local restaurants including Mr. Gatti's and Bruce's Donut King; $3,800 was spent at Wal-mart and $8,000 at the local hardware store. Undocumented checks totaling $60,000 went to one of the sheriff's deputies.

"They say they are fighting the war on drugs," Weiland said. "So they pull people over without reasonable cause, take their money and spend it on donuts."

"They say they are fighting the war on drugs. So they pull people over without reasonable cause, take their money and spend it on donuts."

The investigation prompted another audit by the state's legislative auditor and legislation introduced by the governor that changes burden-of-proof laws and makes it more difficult to seize property.

The computer played a relatively small role in the months-long investigation. However, it helped Weiland and Schwartz take the story beyond anecdotal accounts.

"It placed into context the extent of the misuse," Weiland said. "And it gave us hard evidence to prove how they were using the money."

This story is another example of the benefits to building your own database. Entering hundreds, sometimes thousands, of records can be a daunting task. But you are doing something the agency compiling the records hasn't done, and you are looking at their data in ways they haven't. Secondly, you avoid the "dirty data" issues inherent with data entered by a $5-an-hour clerk.

David Washburn can be reached by e-mail at david.washburn@mcall.com or by phone at (610) 820-6584.

CAR with microdata

By Paul Overberg
USA Today

Microdata is a word from the social sciences. Loosely, it means data at the level at which it's collected. More specifically, it often means the individual records that result from a carefully designed survey where the unit is often a single person, household, incident or something similar. Microdata is usually contrasted with macrodata, a generic term for various levels of summarization and cross-tabulation.

Why bother?

If you're thinking, "Microdata! Fancy word for something I already do!" you're wrong. True microdata can be much more rewarding. You can do cross-tabulations that no one else has ever done – and find news. But it's also trickier. It's often produced by a survey or statistically random selection, so each record represents a larger number of its type in the entire population. But we have to preserve and use each record's weight correctly. Depending on the survey's precision, weights can vary a little or a lot from record to record. It's often collected in a hierarchical file, something akin to interleaving the tables of a relational database. For instance, the Current Population Survey, done monthly by the Census Bureau for the Bureau of Labor Statistics, collects information about 50,000 households into one big file with three record types for their personal, family and household characteristics. Often, a researcher will pull data from the people and household records into a new working file. This is called rectangularizing.

```
Household 1 - all data
       Family 1 - all data
              Person 1 - all data
              Person 2 - all data
              Person 3 - all data
```

becomes in a working file

```
Person 1 - some data and
       Household 1 - some data
Person 2 - some data and
       Household 1 - some data
Person 3 - some data and
       Household 1 - some data
```

This turns the hierarchical file into a flat (rectangular) file that's smaller and easier to analyze. If you've ever analyzed the Fatal Accident Reporting System (FARS) data from the National Highway Traffic Safety Administration, you've faced this. For each fatal accident record, there's one or two or more related vehicle records, each of which has one, two or more related person records.

Before using microdata...

Don't – if you don't have to. Often, summary data is all you need. Start there. Read the standard reports published by the keeper of the data. Look in the back for detailed cross-tabulations if the standard summaries aren't enough. For instance, median household income appears right up front in the Census Bureau's annual report on income. Dig a little and you can find median income by race (all whites). Dig deeper and you can find median income by race by age group (whites 45-54). Dig deeper and you can find median income by race by age by education (whites 45-54 with a bachelor's degree), deeper still for median income by race by age by education by work status (whites 45-54, bachelor's degree, work full-time) and even deeper still for median income by race by age by education by work status by region (whites 45-54, bachelor's degree, work full-time, Midwest).

Talk to whoever keeps the data before you spend a week plowing into a microdata maze. They may have done what you're attempting or have good advice on how to do it (or why not to).

Learn the data thoroughly. This is especially true for microdata based on surveys or hierarchical files. Read the appendix on the sampling design. Read the file list and values. Try to duplicate published calculations.

Learn SPSS or SAS. Either will do, but knowing neither leaves you fumbling with spreadsheets and database managers. They may work, but they're sort of like using a regular srewdriver on a Phillips head screw: You have to work harder and still may get inferior results.

Paul Overberg can be reached at
(703) 276-5427 or by e-mail at
poverberg@usatoday.com

THIS COLUMN IS AN EXCERPT FROM A HANDOUT AT THE INDIANA CAR CONFERENCE.

OVERBERG WRITES OF EXAMPLES OF MICRODATA AVAILABLE ON THE WEB:

- **PUBLIC USE MICRODATA SAMPLE (PUMS) FROM THE DECENNIAL CENSUS OF POPULATION AND HOUSING: WWW.IPUMS.UMN.EDU & WWW.CIESIN.ORG/DATASETS/PUMS/PUMS-HOME.HTML**
- **CURRENT POPULATION SURVEY: WWW.BLS.CENSUS.GOV/CPS/ADS/ADSMAIN.HTM**
- **FEDERAL RESERVE BOARD'S BIENNIAL SURVEY OF CONSUMER FINANCES: WWW.BOG.FRB.FED.US/PUBS/OSS/OSS2/SCFINDEX.HTML**
- **BUREAU OF LABOR STATISTICS SURVEY OF CONSUMER SPENDING: WWW.BLS.GOV/CSXSTND.HTM**
- **CENTERS FOR DISEASE CONTROL & PREVENTION BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM: WWW.CDC.GOV/NCCDPHP/BRFSS**

Pulitzer profiles

By Nora Paul
Poynter Institute

Welcome to the "On the Internet" Pulitzer Prize column.

We'll look at each of the prize winners and point you to versions of their stories (if available on the Web) and related Web sites that might help you if you are covering a similar type of story.

Public Service: *Grand Forks Herald*

Coverage of the flood, blizzard and fire that devastated much of the city, including the newspaper plant.

Read about the win: http://www.gfherald.com/news/daily/415/415front.htm

Read some of the stories: http://newslibrary.infi.net/

The Infinet NewsLibrary site has the archive of *Grand Forks Herald* stories back to 1994. Searching is free, but if you want to read an article, it's $1.00 each.

If you're covering a similar story: Disaster Resources

• Federal Emergency Management Agency: http://www.fema.gov

• EmergencyNetNews: http://www.emergency.com/ennday.htm – a resource of articles, reports and backgrounders on every kind of disaster and emergency, worldwide.

• Disaster Finder: http://ltpwww.gsfc.nasa.gov/ndrd/disaster/ – "A complete index to the best disaster Web sites on the Internet. Period." Well, sort of. Good subject organization of the links but little annotation of the information to be found.

• Disaster Relief: Worldwide Disaster Aid and Information Via the Internet: http://www.disasterrelief.org/Links/ - great compilation of resources on the Internet regarding disasters. Links to Red Cross sites, disaster relief agencies, and reference sites.

• Yahoo's Disaster Links: http://www.yahoo.com/Society_and_Culture/Environment_and_Nature/Disasters/ – links to the following disaster categories: airplane accidents, avalanches, blizzards, Chernobyl, drought in the Midwest, Dust Bowl, earthquakes, fires, floods, hurricanes, landslides, nuclear disasters, oil spills, organizations, tornadoes, typhoons, volcanoes.

Breaking News Reporting: *Los Angeles Times*

Coverage of a botched bank robbery and subsequent police shootout in North Hollywood.

Read about the win: http://www.latimes.com/HOME/NEWS/REPORTS/PULITZER_98/

See and hear the newsroom's reaction: http://www.latimes.com/MULTIMEDIA/VIDEO/shootout.980415.mov

Read some of the stories: http://www.latimes.com/HOME/NEWS/REPORTS/PULITZER_98/N_HOLLYWOOD/

If you're covering a similar story: Crime Resources

• Copnet: http://police.sas.ab.ca/ – great list of links to resources and to city and state police sites.

• The Police Officer's Internet Directory: http://www.officer.com/

• Armed Robbery Page: http://www.ior.com/~jdmoore/ – interesting little page by an entrepreneur wanting to train employees in high-risk jobs about how to deal with an armed robbery. Based out of Spokane, Wash.

Investigative Reporting: *Baltimore Sun*

Series on the international shipbreaking industry, what happens to workers and the environment when discarded ships are dismantled.

Read about the win: http://www.sunspot.net/archive/search/ – You'll have to use their archive of stories going back to 1990 to read about the win – and it will cost you $1.95.

Read the stories: http://www.sunspot.net/news/special/shipbreakers/ – They have packaged the series on their Web site, with photos.

Explanatory Reporting: *Chicago Tribune*

Profile of the Human Genome Diversity Project

Read about the win: http://www.chicago.tribune.com/news/metro/chicago/article/0,1051,ART-7040,00.html

Read the stories: http://www.chicago.tribune.com/news/nationworld/article/0,1051,ART-6989,00.html

Continued on page eleven

OTHER WEB SITES FOR JOURNALISM AWARD WINNERS:

• **IRE MEDALS AND CERTIFICATES:** WWW.IRE.ORG/RESOURCES/CONTEST/INDEX.HTML

• **PULITZER PRIZES:** WWW.PULITZER.ORG/YEAR/1998

• **SOCIETY OF PROFESSIONAL JOURNALISTS SIGMA DELTA CHI AWARDS:** WWW.SPJ.ORG/SDXAWARDS97/1997WINNERS.HTM

• **AMERICAN SOCIETY OF NEWSPAPER EDITORS:** WWW.ASNE.ORG/KIOSK/WRITINGAWARDS/98WRITINGAWARDS.HTM

Continued from page ten:

Winners on the Web

If you're covering a similar story: Human Genome Diversity Project, genetics
• Human Genome Project Information: http://www.ornl.gov/TechResources/Human_Genome/home.html
• Human Genome Program information from the D.O.E.: http://www.er.doe.gov/production/ober/hug_top.html
• The Human Genome Organisation: http://hugo.gdb.org/
• The National Human Genome Research Institute: http://www.nhgri.nih.gov/

Beat Reporting: *New York Times*

Coverage of the U. S. Supreme Court
Read about the win and click to the stories: http://www.nytimes.com/library/national/041498greenhouse.html
If you're covering a similar story: U.S. Supreme Court
• Supreme Court Decisions: http://supct.law.cornell.edu/supct/
• Rules of the Supreme Court of the United States: http://www.law.cornell.edu/rules/supct/overview.html

National Reporting: *Dayton Daily News*

Disclosure of flaws and mismanagement in the military health care system
Read the stories: http://www.ActiveDayton.com/doctors/i/index.html
If you're covering a similar story: Military hospitals
• Department of Defense, Health Affairs: http://www.ha.osd.mil/index.html
• Department of Veteran Affairs: http://www.va.gov
• Listing of Defense Department health care policies: http://www.ha.osd.mil/ppc/num_poli.html – many with links to the documents themselves

International Reporting: New York Times

Series on drug corruption in Mexico
Read about the win and click to the stories: http://www.nytimes.com/library/national/041498mexico.html
If you're covering a similar story: Mexico, DEA, drug smuggling
• Drug Enforcement Administration: http://www.usdoj.gov/dea/
• Office of National Drug Control Policy: http://www.whitehousedrugpolicy.gov/ - use "search" and find items related to Mexico.
• United Nations International Drug Control Programme: www.undcp.org/
• Media Awareness Project: Drugnews Search: http://www.mapinc.org/crime.htm/ – locate articles about drug policy and enforcement from a variety of sources.

Feature Writing: *St. Petersburg Times*

Portrait of a mother and two daughters murdered during a Florida vacation and the investigation of the murders.
Read the stories: http://www.sptimes.com/Pulitzer98/awards.html
If you're covering a similar story: murder
• Webgator: http://www.inil.com/users/dguss/wgator.htm – when you're doing an investigation, these are great investigative resources.
• The Crime Library: http://www.crimelibrary.com/ – links and text to other famous / infamous murder cases. The sources of the reports seem credible, but none are crime writers.

Commentary: *New York Daily News*

Coverage of a Haitian immigrant brutalized by police officers at a Brooklyn station house.
Read about the win and click to the stories: http://search.mostnewyork.com/most/manual/news/mcalary/machome.htm
If you're covering a similar story: police brutality
• The Police Complaint Center: http://www.policeabuse.com/ – a non-profit seeking to compile incidences of police abuse. Located in Tallahassee, Fla.
• Yahoo's Police Brutality links: http://www.yahoo.com/Society_and_Culture/Crime/Crimes/Police_Brutality/ – Most of the sites listed are decidedly biased Web sites. Use with caution, but the sites might be interesting sources for personal stories.

Nora Paul can be reached by e-mail at npaul@poynter.org

• **NATIONAL HEADLINER AWARDS: WWW.PRESSPLUS.COM/HEADLINER/HEADLINER.HTML**

• **GEORGE POLK AWARDS: WWW.LIU.EDU/CWIS/BKLYN/POLK/98AWARDS/HTML#TOP**

• **GEORGE FOSTER PEABODY AWARDS: WWW.GRADY.UGA.EDU/PEABODY**

IF YOU WOULD LIKE TO SEE A PARTICULAR INTERNET TOPIC DISCUSSED IN THIS COLUMN, SEND AN E-MAIL TO BRENT JOHNSON AT BJOHNSON@NICAR.ORG

HANDOUT OF THE MONTH

Steps to data freedom

By Michael J. Berens
Chicago Tribune

Ancient warrior wisdom is based on the belief that the battle is won or lost before the first sword is drawn. Likewise, the fight for data must be plotted before the first request is filed. Understanding the buzz words – the warrior language, if you will – will go far in negotiations.

Study the file layout

Skip this vital first step at your peril. Trust me. Take time to do this or risk major setbacks in time and money. The file layout is a document that shows data organization and structure. It serves as a blueprint for obtaining, importing and analyzing data. Code keys are a list of standardized abbreviations, acronyms or numbers that represent information. For example, states may be represented as numbers in the database.

Study file layouts before making requests. For example, the database layout of the National Highway Traffic Safety Administration lists hundreds of fields involving fatal vehicle crashes. One field, called the "related factors" field, uses a code number to designate if the fatality was related to a police pursuit. An entire series can – and has – been built on one field.

Layouts also reveal a major pitfall for the unwary: packed fields.

Typically, packed fields hold information that is encrypted and condensed as a space-saving technique. Unless fields are unpacked, you may not be able to read the information.

Often, government data is stored in multiple files, called a relational database. Examine layouts for a common field, which is necessary for linking files.

Handling Formats

This is the step where some bureaucrats love to leave reporters in the dust. Learn about the types of formats your software can handle. However, most any database can be copied as an ASCII file. This is an acronym for standardized computer format that is recognizable by all mainstream software (FoxPro, Access, Excel,File Maker Pro, etc.)

Most software also can handle multiple formats, such as DBF or text format or formats managed by software applications.

Next, you must determine how the data is delimited – how it will be separated. Many reporters will ask for data in ASCII format (an industry standard) and delimited by commas. During importing, the computer will separate the data into new fields at each comma. You can have the data delimited by spaces, or tabs, or most anything, if preferred.

Medium messengers

You've followed the first two steps flawlessly, the agency quickly delivers your data – and it's on a 4mm tape cartridge. This happened at the *Tribune* when a reporter forgot to request the medium, which is the all-important third step. Needless to say, few newspapers have 4mm tape drives to read the data. Determine what formats your paper can handle, such as disks or cartridges or tapes. Besides disks, here are some other choices:

- Nine-track tapes – Used by many government agencies, but you will need a special nine-track reader attached to your computer.
- JAZ Cartridges – These tapes hold 1 gig or more of data. JAZ drives are common and increasingly more inexpensive.
- ZIP Cartridges – Similar to JAZ, but typically these cartridges hold less information.
- FTP (File Transfer Protocol) – This isn't a storage medium, but it's a dream way to obtain data. In short, the agency temporarily places the requested data on its Web page, allowing you to transfer the data via the Web into your computer. No tapes, no disks, no special readers. And it's more efficient for agencies.

Roadblocks and solutions

The pursuit of data is no laughing matter, but government response can border on the ridiculous. Here are actual government objections encountered nationwide. The successful strategy centers on getting past the flak phalanx and talking directly to the agency's data processing employees.

- *Our computer can't do that.*
- *You'll never understand the data. It's far too complex. We're afraid you'll make mistakes*

Continued on page thirteen

MICHAEL BERENS CONTRIBUTED THIS HANDOUT AT THE INDIANA CAR CONFERENCE IN MARCH.

TO ORDER A COPY OF ANY HANDOUT APPEARING IN THE PAGES OF *UPLINK*, CONTACT THE RESOURCE CENTER AT (573) 882-3364.

FOR A LISTING OF AVAILABLE HANDOUTS AND TIP SHEETS, BROWSE OUR WEB SITE AT WWW.IRE.ORG/RESOURCES/CENTER/HANDSEARCH.HTML

NICAR'S DATABASE LIBRARY CAN CONVERT DATA FROM ALL TYPES OF GOVERNMENT SYSTEMS. THE STAFF CAN TAKE 9-TRACK TAPES, 4MM DAT TAPES, 3480 CARTRIDGES, PRINT IMAGE FILES OR TRI/TR2 TAPES, AND PUT THEM ONTO CD-ROM IN THE DATABASE FORMAT OF YOUR CHOICE. STAFF CLEANS AND CONVERTS THE DATA, AND PERFORMS INTEGRITY CHECKS. PRICES ARE BASED ON THE SIZE OF YOUR NEWS RGANIZATION. PLEASE CALL THE DATABASE LIBRARY FOR MORE INFORMATION, 573-882-0684.

Continued from page twelve:

Data battle plan

and draw wrong conclusions.

• It's too big. You'll never have a computer to handle it.

• Our software is proprietary, so you can't have the data.

Data negotiations are fought with knowledge and confidence. More than once I've had a new term thrown at me during a meeting. Don't panic, don't flinch. Keep smiling, say no problem, then later run to someone in the newspaper's data processing department to find out what the term meant – and whether it's actually going to be a problem. Don't let the agency bluff you, but don't be afraid to do a little bluffing, too.

Demand specifics, then demand proof. IRE and NICAR have excellent resources for those confused by techno-babble such as megs and gigs (terms related to data storage).

Don't be afraid to ask basic questions. Ask the agency computer people how they relate multiple files: you may learn their method is more complex – but more accurate – than your plan.

Another psychological negotiating ace is to assure the agency that you plan to provide them a copy of all analysis (we're talking aggregate conclusions and totals) prior to publication. I love the moment I get to show the agency what their own data reveals. And more than once, I've been saved from publishing incomplete conclusions. Remember, the potential for errors is exponentially higher in CAR reporting. You're just one keystroke away from an error, always.

Michael Berens can be reached at
(312) 222-4798 or by e-mail at
mberens@tribune.com

Continued from page five:

Excuse-proof data

tion rate for a DUI declined as the number of officers involved increased. That finding made it impossible for police to claim that extra officers ensured a conviction.

We created another database from the logs of the breathalyzer machines. We punched in the names of defendants whose breath-alcohol readings were below the legal limit, then merged the defendant names against the DUI court cases. More than half of those cases had been dismissed outright, meaning that hundreds of people were being charged with DUI even though tests showed they were not legally intoxicated. Although it's hard to prove that police are intentionally arresting innocent people to drum up court overtime, our finding from the data certainly hinted that drivers were being unfairly victimized.

Matching payments

Only one of the departments we examined kept electronic payroll records that matched pay with specific cases. For the other departments, reporters pulled paper files and read time slips.

I-team reporter Lisa Getter, using Microsoft Access, did much of the matching of the overtime database with databases of arrests and off-duty police overtime jobs. We found officers in multiple places at the same time. One year, there were 85 cases of officers who were in court for less than 10 minutes but still received the union-specified minimum of three hours overtime pay.

To detail how officers "piggybacked" onto cases, we built a supplemental database for one subset of cases. Using paper records, the reporters identified exactly which officers performed which tasks on each case. Using an Access front end, reporters entered the case number or defendant name. The data-entry form then automatically put the officers' names into drop-down boxes to fill in each task on a case. That way reporters could point-and-click very quickly and avoid typos.

The computer analysis was lengthy and challenging. But it was rewarding. The largest local police union, which was placing blame elsewhere and claiming officers were being victimized, never even attempted to deny or contradict our findings. In police memos before our series, the problem was downplayed as isolated or exaggerated. The data proved them wrong.

Dan Keating can be reached at
(305) 376-3476 or by e-mail at
dkeating@herald.com

ONLINE CAR PROJECTS:

TO VIEW A LISTING OF LINKS OF RECENT COMPUTER-ASSISTED REPORTING STORIES POSTED ON THE WEB, POINT YOUR INTERNET BROWSER TO WWW.IRE.ORG/RESOURCES/CONFERENCES/TRAINING/CARPROJECTS.HTML

THE SITE INCLUDES A DESCRIPTION OF THE STORIES AS WELL AS LINKS TO IRE AWARD WINNERS. IF YOU WOULD LIKE TO SEE A STORY ADDED TO THE LIST OF LINKS, SEND JACK DOLAN AN E-MAIL AT JACK @NICAR.ORG.

TECH TIP

Cold Fusion

By Ray Robinson
The Virginian-Pilot
and **Tom Boyer**
The Seattle Times

Ray Robinson attended the NICAR advanced boot camp in May 1997.

Not long ago, we all dreamed of the day when everyone in our newsrooms would be able to fire up Access or FoxPro, craft a simple query and retrieve a vital nugget of information on deadline.

Cold Fusion is one of a growing class of software products known as "middleware." On a computer that's been set up as a Web server, it takes queries submitted from a user's Web browser, submits them to a database, places the results in HTML and sends them back to the browser.

Well, it's not going to happen, is it? It's not that journalists can't learn these skills. But unless they're using them every day, the skills won't be there when they need them. If we're really going to make the data easily accessible to them, we need another solution.

That's where Cold Fusion comes in.

Cold Fusion is one of a growing class of software products known as "middleware." On a computer that's been set up as a Web server, it takes queries submitted from a user's Web browser, submits them to a database, places the results in HTML and sends them back to the browser.

The technical advantages of such a setup are too numerous to list here, but a few are worth mentioning:

• Reporters on deadline use Netscape or Internet Explorer instead of Access or FoxPro to retrieve data.

• All the heavy lifting is done on the Web server, not the user's computer. So the person using a 486 with 8 megs of RAM can query a 500,000 record database. All they need is a browser.

• Even Mac users can query the database as long as they have a browser.

If you're comfortable with databases and have a passing knowledge of HTML, you can use Cold Fusion.

The software operates from "templates," which are HTML-like pages with bits and pieces of CFML (Cold Fusion Markup Language) embedded in them. To put it simply, the HTML controls the formatting and style of the page the user sees. The CFML dictates what is retrieved from the database and how it's displayed.

Script studies

As an example, let's say you have a database of names and addresses you want to put on the Web. And rather than putting each record into HTML and displaying it, you want to let users query the database for individual names.

The following code would create a text box where the user could enter a person's last name:

```
<FORM ACTION="RESULTS.CFM"
METHOD="POST">
<b>Enter last name here:</b>
<INPUT TYPE="text"
NAME="Last_name" SIZE="50"
MAXLENGTH="40">
<INPUT TYPE="submit"
value="submit"></CENTER></FORM>
```

When the user submits a last name, the Web server kicks off the Cold Fusion script ("RESULTS.CFM"), which processes the query and returns the information.

Cold Fusion takes the information off the form and runs the query, with the following code from "RESULTS.CFM":

```
<CFQUERY NAME="GetResults"
DATASOURCE="MYDATABASE">
SELECT * FROM NAMELIST WHERE NAME
= #Last_Name# order by Name
</CFQUERY>
```

Continued on page fifteen

COLD FUSION IS MANUFACTURED BY THE ALLAIRE CORP. OF CAMBRIDGE, MASS. GO TO THE COMPANY'S WEB SITE AT HTTP://WWW.ALLAIRE.COM, WHERE YOU'LL FIND PRODUCT INFORMATION AND SUPPORT FORUMS POPULATED BY COLD FUSION DEVELOPERS.

ANOTHER GREAT SOURCE ON COLD FUSION IS THE BOOK "THE COLD FUSION WEB DATABASE CONSTRUCTION KIT," WRITTEN BY BEN FORTA AND PUBLISHED BY QUE CORP. IT COMES WITH A SINGLE-USER COPY OF COLD FUSION TO INSTALL TO CHECK OUT THE PRODUCT BEFORE INVESTING IN IT.

Continued from page fourteen:

Enhanced access

The query is actually processed by a snippet of computer code called the ODBC (Open Database Connectivity) driver. The results are passed back to Cold Fusion, which sticks them into HTML pages that the server sends back over the network to the user.

That part of the script looks like this:

```
<CFTABLE QUERY="GetResults">
<CFCOL HEADER="Name" WIDTH="20"
TEXT="#Name#">
<CFCOL HEADER="Address" WIDTH="30"
TEXT="#Address#"></CFTABLE>
```

"GetResults" is the name of the query that has just run. The Name and Address inside the pound signs are references to fields in the database table. Users, viewing the results with a browser, would see something like this:

```
Name            Address
BOYER           3022 NW 66th
ROBINSON        P.O. Box 449
```

Like many other Web middleware products, Cold Fusion talks to databases through ODBC, a database query standard. That's good. ODBC drivers can access a variety of data formats on a local machine, including Access, dBase/FoxPro, Excel and Paradox. ODBC drivers can also access higher end client-server databases such as Microsoft SQL Server, Oracle or IBM's DB2.

The bad news is that the speed and quality of the drivers varies greatly. Older Access "Jet" drivers can be slow and prone to crash, for example. The Visual FoxPro driver you can download from Microsoft is much faster than the FoxPro driver that generally comes with ODBC packages.

Enabling middleware

So what do you need in the way of hardware to put Cold Fusion to work?

You need a Web server. But that doesn't necessarily mean a monster machine with 128 megs of RAM and an 8-gigabyte drive.

In fact, a Web server is nothing more than a PC with Web server software – such as Website or Microsoft Internet Information Server – installed on it.

If you want to start out small, you can probably get away with setting up one of the better PCs on your network as a Web server.

At *The Virginian-Pilot*, we test Web database applications on a Pentium 200 with 32 megs of RAM. The setup has so far worked fine, even with more than one user accessing the data at the same time.

After testing, we move the applications to a bigger server maintained by our online edition for two reasons: (1) It gets them outside the corporate fire wall where our bureaus, which still access the Web through dial-up connections, can get to the data; and (2) It allows us to offer the general public access to some of the data.

But for those considerations, we could probably get by just fine running them on a makeshift Web server in our newsroom.

Keep in mind, though, that more RAM is always better. And with 32 megs of RAM now selling for about $50, don't short your machine on memory unless you have to. Extra RAM also saves wear and tear on more expensive hard drives.

And give your server the best hard drive you can afford. It will dramatically increase performance. Look into an Ultra Wide SCSI hard drive running at a high RPM (7,200 to 10,000). Probably the best available for Web servers right now is the Seagate Cheetah series.

When you begin to develop Web database applications, you'll find yourself confronted with a lot of products with exotic – and baffling – names like Active Server Pages, Visual Interdev, Cold Fusion and countless others.

Some Web site building software, such as Microsoft Front Page 98, comes with tools for building dynamic applications via ODBC connections to databases.

The growing list of products makes for a pretty confusing picture. We can't say that Cold Fusion is the best solution, but we can say that it has worked well for us at *The Virginian-Pilot* and *The Seattle Times*.

You can waste a lot of time searching for the perfect product. Or you can pick one that you're comfortable with, learn it, stick with it – and get your data on the Web. That's what we did with Cold Fusion.

Ray Robinson can be reached by e-mail at rsquared@pilotonline.com
Tom Boyer can be reached by e-mail at tboy-new@seatimes.com

IF YOU'D LIKE TO SEE COLD FUSION IN ACTION, CHECK OUT THE PILOT'S DATABASE OF CAMPAIGN CONTRIBUTIONS TO STATE ELECTED OFFICIALS AT HTTP://WWW.PILOTONLINE.COM/VOTER. FOLLOW THE LINKS UNDER "SEARCH CAMPAIGN CONTRIBUTIONS."

Bits, Bytes and Barks

Indiana CAR Audio Tapes

Audio tapes from the sessions at Indiana CAR are now available for ordering. For more information or an order form, point your browser to *www.ire.org/resources/nicar/conferences/indiana/audio.html*

Mailing Lists

IRE maintains a number of mailing lists, some with longstanding presence and some just getting started.

To subscribe to **IRE-L** or **NICAR-L**, send an e-mail to listproc@lists.missouri.edu. In the body of the message, type: subscribe NICAR-L <yourname>
subscribe IRE-L <your name>

To subscribe to **CFIC-L**, the newly formed Campaign Finance Information Center listserv, send an e-mail to majordomo@campaignfinance.org. In the body of the message, type: subscribe CFIC-L <your e-mail address>.

To subscribe to the **INTRACAR** mailing list, which discusses running and developing Intranets, send an e-mail to majordomo@nicar.org. In the body of the message, type: subscribe intracar <youre-mailaddress>.

Also, check out the IRE-L and NICAR-L mailing list archives on our Web site at *www.ire.org/resources/ire-l/index.html* and *www.nicar.org/resources/nicar/nicarl.html*. You can see posts to both lists organized by thread, author and date. The list archives are available in *.html or plain text format.

Data Library Additions

NICAR's database library has recently added two databases to its collection. The IRS Business Master File database contains a listing of exempt organizations. Some fields include the organization name, address, exempt classification category, primary function of the organization, asset amount and income amount.

As well, the library has acquired the 1996 FBI Uniform Crime Report databases: six separate databases gathered from law enforcement agencies around the country. They are arranged by reporting agency and broken down by month. The library can be reached at (573) 884-7332.

Campaign Finance Information Center

Campaign finance data from thirteen states is freely downloadable from the CFIC at www.campaignfinance.org. We also have links to ten online search engines hosted by other non-profits and state boards of election. We will soon release our "universal" online search engine so you can type in a contributor from your state and see where else they are giving.

The more inclusive this database, the better for everyone. So if you have state or local campaign data you want included, contact CFIC Coordinator Jack Dolan at jack@nicar.org or (573) 882-1982. The CFIC credits everyone who contributes data.

Investigative Reporters and Editors, Inc.
138 Neff Annex
University of Missouri
School of Journalism
Columbia, MO 65211

NON-PROFIT ORG.
U.S. POSTAGE
PAID
PERMIT NO.286
COLUMBIA, MO. 65211