February 1998

Uplink

A newsletter for the National Institute for Computer-Assisted Reporting

INSPECTION DU JOUR

Restaurants Revealed

By Joel Grover
KCBS-TV

When I first proposed the idea of a series on restaurant conditions, it seemed like such a basic investigation. I knew similar investigations had been done for years in other cities. I would have never thought a series like this would attract a firestorm of attention, would be the subject of dozens of newspaper articles, and result in major changes in the law. But that's just what happened when KCBS-TV aired "Behind the Kitchen Door" last November.

(For related stories, see Stuart Watson's restaurant inspection reprise on page two and a review of a Cox News Service project on food processing plants on page three.)

The story began with a tip: a former L.A. County health inspector told me the county was doing nothing to force filthy restaurants to clean up. A quick computer search showed me that no other local media had ever investigated L.A. County's restaurant inspection program, so it seemed like a good subject to probe.

I began my research with what seemed like a simple request to the L.A. County Health Department for its computer database of restaurant inspections. At first the county said such a database didn't exist and later said certain restaurant data wasn't public. Fortunately, my station was willing to fight for this data. CBS lawyers battled L.A. County back and forth for 8 months. Only when we told them we were taking them to court did the health department hand over its database.

Hired help

KCBS hired NICAR to crunch the data. Clearly we could not have done this ourselves. L.A. County's database contained hundreds of thousands of records on 20,000 restaurants. The county tried to send us the data in a form that would have made it virtually unusable. One phone call from NICAR to the health department solved the problem.

Unsettling scores

Our producer, Adam Symson, and I brainstormed with Andy Lehren and Jack Dolan at NICAR about all the ways we could crunch the data. We used Foxpro 2.6. We wanted to know how often restaurants were inspected. We wanted to see what percentage of restaurants were averaging poor to failing

Continued on page four

FREE BOAT TO CHINA

All Expenses Paid

By Kara G. Morrison
NICAR

After Congress tightened rules on member travel funded by private groups in late 1995, Gannett News Service decided to check if the new rules had made a difference.

"We wanted to assess how much travel members of Congress and their staffs have taken from private groups with a stake in Capitol Hill legislation," said GNS special projects editor David Milliron.

The result was a GNS special report titled "Congress: Have invitation, will travel."

The special report, which ran predominantly in many of Gannett's 91 daily newspapers in November, included an overall computer analysis and story on congressional travel, along with stories detailing which special interests paid the highest travel tabs, which destinations were most common, which committees

Continued on page five

CHECK, PLEASE

Update

Check, please. This food is not so glorious.

On this month's menu, *Uplink* serves up 3 CAR projects exposing deficient food inspections.

Joel Grover recounts a KCBS series on L.A. restaurant inspections, and Stuart Watson details a WRAL series on North Carolina's misleading ratings system. As a third course, a summary of a Cox News Service look into sparse USDA punishments of processing plants is also available. And Nora Paul tops it off with a survey of relevant links.

A review of a Gannett News Service analysis of privately sponsored congressional travel and Neil Reisner's advice on doing short-term CAR projects on breaking news are also inside.

Plus, there's the goings-on at a recent NICAR advanced seminar and a workshop on health care.

PAGE 10
First Ventures
PAGE 11
On The Internet
PAGE 12
Stats from the road
PAGE 13
Tech Tip
PAGE 16
IRE/NICAR Notes

Uplink

February 1998
Volume 10, Number 1
A newsletter for the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston
brant@nicar.org

MANAGING EDITOR
Brent Johnson
bjohnson@nicar.org

SENIOR CONTRIBUTING EDITORS
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
Margaret Murphey
Curt Poff
Debbie Roberts
John Sullivan
Alistair White
Paul Walmsley
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.

Behind the Ratings

By Stuart Watson
WRAL-TV

Most of us eat out hundreds of times a year – sometimes more often than we eat at home. While some CAR veterans look down their noses at reports on restaurant health inspections as elementary or overdone, few topics strike closer to home with the public. Many newspapers have long posted weekly sanitation inspections as part and parcel of the news of record. Now, databases of restaurant inspection reports are ready-made for the Web. They give the public fingertip access to raw information about local eateries.

WRAL's series began when a customer at a McDonald's restaurant in Raleigh, N.C., took a close look at its sanitation inspection report posted on the wall. Surprised, she called us to say that, while the local health department had awarded the McDonald's an "A" sanitation grade, the inspector had noted maggots growing in an unused ice cream machine.

Making the grade

So what does an "A" rating mean, anyway?

We filed an Open Records Act request to get our county's database of sanitation inspection reports to find out. Rules, however, vary from state to state. As we say in TV, check your local listings.

North Carolina, South Carolina and Nevada are the only states that use A-B-C ratings. Most states still use variations of a 100-point scale. Some are moving toward the Food and Drug Administration's latest "model" food inspection rating system, which calls for a satisfactory/unsatisfactory approach that is obviously less conducive to computer-assisted reporting.

North Carolina's database (including only one fourth of all counties) includes fields for type of violation, numbers of points deducted from a 100-point scale, final score and – most valuable of all – the inspector's comments. We found restaurants reportedly having maggots, raw sewage backups, and rats breeding in the dumpster that still received "A" ratings.

Inspectors routinely cited restaurants for less sensational – though more hazardous – violations like food served at improper temperatures and raw meat contaminating cooked food, yet gave them "A" ratings as well.

The inspectors aren't to blame.

The state's 100-point system allows a maximum deduction of only five points for each type of violation – not each individual violation. So a restaurant with a series of relatively innocuous but different violations could score lower than a restaurant with a series of threatening but similar ones.

As is the case with almost all computer-assisted reporting, the database was not the final story – only the beginning of questions.

We adopted the methodology used by Mike Wendland at WDIV-TV in Detroit: culling the lowest and highest scores, then looking at files at the health department to build a much smaller spreadsheet with at least 3 inspections for the group of low-scoring restaurants.

Expecting surprise

This removed the "just-had-a-bad-day" defense and showed restaurants that scored poorly over time.

It also contained a surprise. Seven of the 10 lowest scoring restaurants posted "A" ratings. Because inspectors are required to re-inspect within 15 days, some restaurants would routinely score low on surprise inspections – then clean up their act. The lower and arguably more representative grade from the surprise inspection hangs in public view for two weeks tops; the higher grade from the not-so-surprise inspection hangs on the wall for public consumption for months until it's time for another surprise inspection.

Our report aired just as the state was debating changing from letter grades to number grades. So we were able to follow the debate with new information and context.

Our experience is by no means unique. But in this instance, just because a topic has been covered well in another market does not mean your audience would not benefit from another treatment of it locally – especially if you can augment the on-air story with relevant information online.

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

Processor violations

By Brent Johnson
NICAR

Hundreds of meat and poultry plants operating throughout the country are repeatedly issued citations following inspections by the U.S. Department of Agriculture. But the enforcement record against these gutting operations has most often been gutless. An investigative story by Elliot Jaspin and Scott Montgomery of Cox News Service found that few plants – even ones cited hundreds of times – are ever forced shut by the USDA.

Critical condition

Jaspin and Montgomery used FoxPro to analyze the massive "Feedback File" database provided by the USDA following a FOIA request. Although they requested files from 1990-1996, the USDA only released information from 1996.

"When you have large databases, the whole idea is to divide and conquer, to be able to work in a feasible fashion," Jaspin said.

The reporters wanted to know how many "critical violations" – defined by the USDA as a plant condition "certain" to cause both food contamination and to reach and then harm consumers – occurred at given plants. "We were expecting 200 tops for a particular plant," Jaspin said. However, they found a poultry plant in Arkansas operated by Tyson Foods that had amassed 1,700 violations. That's when they knew they had a story.

File me up, file me down

The USDA database arrived on CD-ROM organized by a series of subdivisions. The data were divided into quarters of the year. Each quarter was divided into 26 regions. Each region was divided by week. "So there were 13 zip files for each region, and within those, when unzipped, were different 'circuits,'" or strings of meat and poultry plants, Jaspin said. "We're talking in the neighborhood of 5,000-6,000 little databases."

Transferring the data from the CD-ROM to the hard drive using a batch file presented the first challenge. Then appending the zip files into databases for each quarter awaited. "We had to smash (the unzipped files) into one large database – actually four. One database would've been too unwieldy – over 25 million records."

Divide and conquer

Dirty data always lurks. Jaspin and Montgomery had to work around an identification problem. They thought each plant would have a unique number but then found individual plants entered multiple times. By cleaning the unique number problem up, they could link violations from inspection records to particular plants by joining tables.

Finally ready to analyze the data, they created "subset databases" to answer particular questions, such as the frequency of critical violations occurring at plants.

"When you have large databases, the whole idea is to divide and conquer, to be able to work in a feasible fashion," Jaspin said.

Jaspin and Montgomery also analyzed the data by geographic distribution of violations and plant size. Smaller databases were created and sent to individual Cox newspapers so they could pursue local stories.

USDA puts up, but doesn't shut down

Through their database analysis of 1996 inspections, the reporters discovered 299 plants with 100 or more "critical violations" and 7 with more than 1,000 . That year, the USDA shut down only 6 plants.

"Heavens," an epidemiologist at the Centers for Disease Control and Prevention told the reporters upon hearing those numbers. "The burning question is why is a plant with that many violations – why has the (federal) inspector not been pulled from that plant? And I have no idea."

Brent Johnson can be reached at
(573) 882-0684 or by e-mail at
bjohnson@nicar.org

NICAR IS IN THE PROCESS OF ACQUIRING THE SUMMARY DATA OF THE USDA'S PROCESS DEFICIENCY REPORTS, WHICH INCLUDE A LIST OF VIOLATIONS – CRITICAL, MAJOR OR MINOR – BY PLANT.

A PROFILE OF PAUL ADRIAN, AN INVESTIGATIVE REPORTER AT WBNS-TV IN COLUMBUS, OHIO, WHO HAS ALSO USED THE SUMMARY DATA IN A STORY, WILL APPEAR IN NEXT MONTH'S ISSUE.

Continued from page one:

Dirty kitchens

scores on their inspections. And if restaurants were failing every inspection, were they penalized or shut down?

Our analyis revealed that L.A. County's restaurant inspection system was not doing its job of protecting the public health. We found that many restaurants hadn't been inspected in over 2 years, even though the health department had previously insisted every place gets at least 3 inspections a year. We found that nearly 30 percent of L.A. area restaurants averaged poor to failing scores on health inspections. The data showed that most of these places were never penalized or closed down.

Finally, NICAR ranked all 20,000 restaurants from the lowest to highest scoring. We were amazed when we examined that list. Some of the most respected and popular restaurants in Los Angeles were also getting the worst inspection scores.

Two days after our first report, the Los Angeles County Board of Supervisors publicly grilled health department officials about the problems KCBS uncovered. The department then started cracking down.

Going undercover

We then decided to conduct an undercover investigation. Using the list that NICAR produced for us, we chose several low-scoring restaurants and sent researchers to work there.

Our computer analysis and our undercover tapes gave us enough material to do many stories. What was originally intended to be a 2-part series turned into an 11-part series. Our first story exposed filthy conditions at many popular L.A. restaurants and used facts from our computer analysis to show the problem was widespread. On the second day we exposed the failures of the inspection system to keep restaurants clean.

We then did several stories using the rankings of restaurants by average inspection score. One story focused on the lowest-scoring restaurants in town; another highlighted the cleanest places.

An activated audience

KCBS was deluged with calls from viewers wanting to see the entire list of restaurant scores. We put that list on our Web site, www.Channel 2000.com. In the following weeks, it received an unprecedented one million hits.

Our initial reports generated hundreds of tips, which gave us material for many solid follow-ups. We reported that a grand jury had warned the health department years earlier that there could be an epidemic of food poisoning if it didn't get tougher on unsanitary restaurants. (The County ignored that warning). We also reported how the health department had a secret surplus of millions of dollars that it could've used to hire more restaurant inspectors.

Until we broke that story, health officials had been saying they were short on money to beef up their inspection program.

What amazed us most was the fallout and public outcry from our series. Two days after our first report, the Los Angeles County Board of Supervisors publicly grilled health department officials about the problems KCBS uncovered. The department then started cracking down, closing more than 300 unsanitary restaurants, including one owned by Los Angeles Mayor Richard Riordan. Three weeks later, the board adopted a sweeping overhaul of Los Angeles County's restaurant inspection system. Because of our reports, restaurants must now post letter grades reflecting their latest inspection scores. And restaurant workers will be required to get training and certification in safe food-handling practices. Three months later, we are still doing follow-up stories from what we all thought would be just a basic investigative story.

Joel Grover can be reached at
(213) 460-3284 or by e-mail at
lax@primenet.com

According to Stuart Watson, his station's restaurant inspection series "demonstrated better than any of our previous investigative projects the ability of WRAL Online to supplement on-air reports." He provided Web sites for similar projects at other organizations:

- **WRAL Online**
  http://www.wral-tv.com/news/wral/5investigates/1997/rest-eval-data
- **WDIV-TV Detroit**
  http://www.wdiv.com/dd4.html
- **WSB-TV and The Atlanta Journal-Constitution**
  http://www.accessatlanta.com/local/inspections
- **The Deseret News**
  http://199.104.95.5/misc/rest/restform.htm
- **KCBS-TV Los Angeles**
  http://www.channel2000.com

Continued from page one:

Travel tabs

traveled most frequently, and which Congressional spouses earned the most frequent flyer miles.

The new rules require members of Congress and their staffs to explain how such travel relates to their work and to disclose its sponsor, destination, cost and purpose within 30 days.

Key findings

The GNS database investigation found that private industry and non-profit groups spent nearly $8 million in 16 months on 5,100 trips.

Other key findings included: China was the hottest congressional travel destination; the Nuclear Energy Institute spent $123,000 on congressional travel while lobbying to send radioactive waste to Nevada during 1996-97; the tobacco industry spent $70,000 on 27 trips for congressmen during the same period; and the House Commerce Committee, which oversees telecommunications and electricity deregulation, topped the list of well-traveled congressional committees, taking nearly 200 paid trips in 16 months.

"This is the type of information Congress doesn't want in electronic form because it's too easy to see patterns among trips when it's in a computer database," Milliron said.

GNS chief political correspondent Chuck Raasch and regional correspondent Fredreka Schouten, together with Milliron, spent three months on the project. They collected the paper records on the more than 5,100 trips between January 1996 – when the new reporting requirements kicked in – and early May 1997. The data from the reports was then keypunched by two interns into six Borland Paradox databases.

Milliron said The Center for Responsive Politics subsequently keypunched the 1997 travel records, which is now available on its Web site.

"I was pleased to see they picked up on this. We feel this is an important thing for the public to know," Raasch said.

"This is the type of information Congress doesn't want in electronic form because it's too easy to see patterns among trips when it's in a computer database," Milliron said.

Roadblocks

Milliron explained there were a few pitfalls during the project. Because the majority of the reports were handwritten, GNS had to painstakingly decipher the name of each traveler on the forms using various congressional directories. And many of the travel records identified sponsors not by company names but by acronyms, requiring additional research.

Raasch said members of Congress also did a poor job of filling out the forms. Often, writing was illegible or total expenses (including transportation, lodging, meals) were added incorrectly.

And in one instance, Milliron said, House Speaker Newt Gingrich reported the cost of his trip in Swiss francs, which GNS converted to U.S. dollars using the exchange rate at the dates of his trip.

"If that's not trying to abate the system, I don't know what is," Milliron said.

Included in the database were the trip sponsors, names of travelers, destinations and the cost and dates of travel. GNS created a separate database for travel related to committee work.

Once the analysis was complete for the national package, Milliron worked with the bureau's regional reporters, who wrote about three dozen stories on the travels of elected representatives serving their regions.

Overall, Raasch said they were happy with the play the series received.

"As we tried to point out, there are a lot of new ways or different ways for money to get into this process," Raasch said.

Kara Morisson can be reached by e-mail at kara@nicar.org

THE 1997 CONGRESSIONAL TRAVEL RECORDS CAN BE ACCESSED ON THE WEB SITE OF THE CENTER FOR RESPONSIVE POLITICS AT HTTP://WWW.CRP.ORG/CRPDOCS/9798TRAVEL/TRAV9798.HTM

Demystifying CAR

By Neil Reisner
The Miami Herald

Okay. You've heard one too many stories about the yearlong CAR investigation using umpity-two databases, way-too-sophisticated statistics and a Pentium II 300 with 556 megabytes of RAM and a 200-gigabyte hard drive.

That just ain't the reality you live in.

But you've also heard one too many stories from one too many CAR wizards wearing a pointed cap with stars and crescents on it about the turn-on-a-dime quick hits that they produced merely by blinking their eyes.

They can be done. With just a little bit of forethought, it's possible – even easy – to apply CAR to your everyday work, even to breaking news.

That's what we've been doing at *The Miami Herald's* edition in Broward County, which, despite its connection to a major metro, is really not very different from a medium-sized daily.

That's because we put out our own edition of the paper in Broward with a remade front page, a local news section, bi-weekly neighbors sections, independent calendar listings, local sports – the whole package. We struggle along with the same limitations of staff and resources that confront most other papers. And we compete fiercely with the *Sun-Sentinel* of Fort Lauderdale, with whom we're at war.

Here are four examples of the ways we've used CAR in recent months – not to produce mega-stories but to add value to the routine. (And, for spice, examples of how we could have done better.) None took more than a day, and most were finished in several hours.

Electing school boards

Education is the big political issue here in Broward, where the local Board of Education has about the same credibility as used car dealers and the schools are overcrowded despite the fact that they've spent $2 billion on construction over the last 10 years. The action involved an election last fall deciding whether to scrap the traditional system of electing school board members at large for a system in which most members would represent specific districts.

The change won overwhelming approval, much to the surprise of some pundits (and school board members) who figured what we call the "condo vote", mostly elderly voters loyal to the locally ruling Democrats, would rally to the status quo. The question was: How did this happen?

We covered election night as usual, downloading overall results from a BBS run by the Superintendent of Elections. The day after the election, though, precinct-by-precinct results were available. Usually, they'd been provided on paper, but last year we got them on floppy disk in a tab-delimited ASCII format.

From there, it was a simple matter to import the data using Excel's import wizard and then to move it over to Access. Teaming up with the political reporter, who knows the precincts like the back of his hand, we were able to isolate model precincts and analyze which way the vote went. The result: Young families and minorities came out in force to support a move they deemed important. The condo vote stayed home. With turnout only about 5 percent, it didn't take much to turn the election.

Boating accidents

It was a real tragedy. A powerboat – the kind dope smugglers and men in midlife crisis love – was speeding up the Intracoastal Waterway in the dark of the night when it ran over a small cabin cruiser out for a midnight cruise, just about cutting it in half. Six people died, and more were critically injured – making it the worst boat accident in years and one that we covered intensely.

We already had about ten years' worth of data on recreational boating accidents from the Florida Marine Patrol. The dirty little secret was that it was still in the fixed-field ASCII format it arrived in. That meant writing import specifications for four tables with lots and lots of fields. And on deadline, in a competitive situation, that wasn't practical. (Remember what I said about "a little bit of forethought?" I should've imported it ahead of time.)

NICAR to the rescue. A quick phone call to Missouri and the good folks at NICAR extracted five years' worth of Florida data for

Continued on page seven

TO FIND OUT HOW TO DOUBLE CHECK YOUR CITY'S MATH, 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 six:

Breaking news

us from the national data set they have. And – may God bless the inventors of the Internet – within an hour it was winging its way into my e-mail box.

Once in-house, it became clear that we'd only be able to use two years' worth of data because of changes in the field layout. But, with only a few straightforward queries, we were able to find out how many boating fatalities there had been during that time statewide and along our stretch of the coast, how many accidents, the boats most commonly involved in accidents and the most common cause of accidents. We printed out the information, turned it over to a graphic artist and put the data in the next day's paper.

We could have done better, though. As is too often the case, there was insufficient communication between the data guy (me), the reporter and the editor on the story. That meant the reporter didn't get the data soon enough and was insufficiently briefed on what it meant. Keeping everyone on the same page is real important.

Multiple births

A local woman gave birth to quadruplets a few months ago. And, of course, the hospital PR folks were on the phone with us all agog and, of course, we prepared to go out and do something on the blessed event. (This was before the "Magnificent Seven" in Ohio made mere quadruplets passé.)

Discussing this in the morning news meeting, an editor said words to the effect of: "Geez. Every time quads are born we take the same picture and write the same story. Can we do something different this time? I wonder how much it will cost to raise these kids?"

We went out on the Web and found, oddly enough, a college professor's class lectures noting the USDA's annual cost estimate of child-rearing. A quick trip to the USDA site produced a press release announcing the most recent estimate and a full report on the estimates in *.pdf format. The USDA estimates how much it costs to raise one child. But you can't just multiply that by four to get the cost or raising quads.

A call to the contact named on the press release got us to a USDA economist who told us the formula we needed. Plugged it all into a spreadsheet and handed it off to the graphics guy.

Religious migration

With around 270,000 Jewish residents, Broward County makes up the fourth largest Jewish community in the country. That's why we pay a fair amount of attention to Jewish issues. So, when a long-established synagogue decided to sell its building to a charter school and move a few miles north to a more densely populated area, it was news.

The local Jewish federation had recently released a demographic study of the local community with a table showing the number of Jewish residents in each ZIP code. They faxed it to us, and a few minutes later it was typed into a database table.

We'd recently started playing with mapping software and were fortunate enough to have a copy of ArcView. It didn't take much to hang the ZIP code table onto a base map of ZIP codes we'd gotten from the county. (Forethought worked here – the base map was all ready, completed by my colleague Dan Keating.)

Once we drew the map, we saw clearly that the two ZIP codes surrounding the synagogue were among the densest Jewish areas in the county. And the area to which the congregation was moving was among the least populated.

That let us ask much more focused questions and explain that, even though the synagogue was located in a Jewish area, most of its members didn't live there. The area where the congregation was moving was closer to where most congregants lived and, more importantly, near a neighborhood where young Jewish families were moving.

Were any of these stories prize-winners? Hardly. And, with the possible exception of mapping software, they didn't use any software other than the standard spreadsheets, database managers and web browsers we all have. It's just a matter of thinking about the stories we're writing and how a little bit of data can make them better.

Neil Reisner, formerly NICAR training director, can be reached at (954) 985-4571 or by e-mail at nreisner@web2000.net

BOUND COPIES OF THE 1997 ISSUES OF UPLINK ARE NOW AVAILABLE. CONTACT IRE AT (573) 882-2042. THE COST IS $25 FOR IRE MEMBERS OR $30 FOR NON-MEMBERS, PLUS SHIPPING.

ADVANCED SEMINAR

Client/server solutions

THE ADVANCED SEMINAR DREW A SELLOUT CROWD FROM *THE SAN JOSE MERCURY NEWS*, *USA TODAY*, *DETROIT FREE PRESS*, *BALTIMORE SUN*, *DAILY OKLAHOMAN*, *LEXINGTON HERALD-LEADER*, ASSOCIATED PRESS AND WCCO-TV IN MINNEAPOLIS.

GEORGE LANDAU AND TOM TOROK WILL REPEAT THEIR SEMINAR ON WEB APPLICATIONS WITH NT AND SQL SERVER THIS SUMMER, AUGUST 14-17. FOR MORE INFORMATION OR TO REGISTER, CONTACT KATY FANNING AT (573) 882-0684.

By George Landau
NewsEngin Inc.

Take a dozen experienced CAR jockeys from around the nation, lock them in a NICAR computer lab for three days at the mercy of two hard-core data wonks, and what do you get?

For starters, you get a dozen CAR jockeys who can install and maintain a Web-based, client/server data warehouse for their newsrooms. And you get to brag about the successful inaugural run of NICAR's advanced seminar.

The advanced seminar arose from a basic question that our crowd has been wrestling with for quite a while: how can CAR specialists manage their valuable data so that the newsroom can perform basic queries without bugging the specialist, and perform fancier queries without a lot of logistical hassles and network slowdowns?

The answer is a client/server database manager with a Web interface. For this particular seminar, the teachers – myself and Tom Torok of the *Philadelphia Inquirer* – instructed our colleagues in the 100% Microsoft solution that consists of Windows NT as the operating system, SQL Server as the database manager, and Internet Information Server (IIS) as the Web server. This all-Microsoft lineup is certainly not the only way to go, but Torok and I believe it provides a strong combination of low-cost, high-performance and predictable results. Besides, what journalist really wants to learn UNIX?

Day One

We began with an hour-long discussion of client/server databases and Web technology – why they make sense for sharing data within a newsroom. The idea behind client/server is that the database queries run on a beefy PC (the server), which accepts commands over a network in the form of SQL statements. The server executes the SQL statement and passes only the results back over the network to the user (the client). The client software can be a Web browser or a desktop data tool like Access, Approach, Paradox or Excel. The client software can use a graphical interface, like a Web form or the Access query builder, to generate the SQL statements; your users don't have to know SQL.

Client/server is generally a better solution than simply putting an Access or FoxPro database on a server and letting your users open the database over the network. While this might work for simple queries on indexed fields, for anything more ambitious it means pumping the entire database through the narrow (and typically congested) pipes of your network. This tends to make the network crawl for everyone else in your newsroom, especially the folks in photo who inevitably need to move several large images at the same time your query is scanning a 1-gigabyte table.

So that they could repeat the process in their own newsrooms, we led the participants through the entire drill of installing Windows NT Server, SQL Server, IIS and the numerous bug fixes ("Service Packs") that Microsoft provides subsequent to any product's release. Fortunately, we hit just enough snags in this process to make for an authentic experience, and by day's end each participant was the new administrator of a fully-loaded NT Server.

Day Two

We spent the morning learning about security, configuration and administration in NT Server, SQL Server and IIS. We sailed through the creation of SQL Server devices, databases, tables, views and indexes. We dabbled with various settings and discussed ways to maximize performance.

Once we had a grip on SQL Server, we used ODBC (Open Data Base Connectivity) to connect from Microsoft Access to SQL Server over the network. We transferred data from Access to SQL Server using ODBC alone and in combination with SQL Server's BCP (Bulk Copy Program).

With some data on the server, it was time to install a Web interface. NewsEngin provided all seminar participants with an application called UniQuery and a license to deploy it in their newsrooms. UniQuery's greatest value lies in its ability to perform global searches of all the databases on your server. A user can enter a name, address, birthdate, etc.; regardless of

Continued on page nine

Continued from page eight:

Managing CAR

what those fields are actually called, in the particular tables being searched, the user gets back search hits in the form of hyperlinked data to follow to related or detailed data.

UniQuery is written as an Active Server Pages (ASP) application, which is essentially a version of Visual Basic for Microsoft's Web server. Microsoft has positioned ASP as the language of choice for NT applications (like UniQuery) that do most of their work on the Web server, as opposed to the browser.

Day Three

Using an online workbook he'd written as an ASP application, Torok guided everyone through the key steps in building a Web application that can query or collect data. ASP applications are just Web pages that contain ordinary HTML and Visual Basic Script (the text aspect of VB — the "If...Then," "For...Next" kind of stuff). Among the topics he covered were the use of conditional logic and loops, input forms, the "post" and "get" HTTP actions, input validation, URL redirection, dynamically generated hyperlinks, submit buttons, the "request," "response" and "session" objects, string functions, numeric functions, database querying and database navigation. It was an awesome performance that left folks wishing for a fourth day spend more time applying it all.

Alas, most of us had a plane to catch. But there's always e-mail...

George Landau can be reached by e-mail at george@newsengin.com

HEALTH CARE WORKSHOP

Risky Behavior

By Seth Hemmelgarn
NICAR

When you're talking about things like margins of error, chi-squares and weighted data, it's easy to sound esoteric. But journalists and epidemiologists from around the country transcended the stereotype Jan. 23 – 25 at the first-ever Risky Behavior health care conference. Sponsored by IRE and NICAR and the Centers for Disease Control and Prevention, the workshop gave journalists and epidemiologists the chance to share ideas and concerns face-to-face.

"Workshops like Risky Behavior put beat-specific, analysis-oriented sessions over geekism," said David Heath, director of computer-assisted reporting at the *St. Louis Post-Dispatch*. "I think that they serve to make CAR more relevant in the newsroom."

People on both sides came away from the conference with a new appreciation for each other. "I think the researchers learned just as much about how reporters are using CAR as we learned about the rigors of social science," said Paul Garber, CAR specialist at the *Post-Dispatch*.

The two groups have at least one common goal: giving people accurate information. Both sides use extensive research to do this.

A large part of the seminar was spent discussing the CDC's Behavioral Risk Factor Surveillance System. The survey questions in the system cover topics ranging from seatbelt use to AIDS testing. Answers are compiled in a database that the CDC makes available.

Sarah Cohen, NICAR training director, taught participants how to get statistics from the data using SPSS, a statistics program that lets users compile statistics with the push of a button.

"The statistics-speak was pretty understandable," said Diane Targovnik, news editor of *Managed Health Care News*. "All the speakers emphasized that you can't look at data and run to a conclusion."

One researcher clarified for Heath that with enough interest and perseverance, anyone can do fairly sophisticated analysis. "Not that I expect many reporters to start doing double-blind trials," Heath said.

The conference also included information on other types of health-care data, where to get it and how to use it. A Web site with follow-up information is in the works.

Seth Hemmelgarn can be reached by e-mail at seth@nicar.org

RISKY BEHAVIOR ATTENDEES CAME FROM *THE CHICAGO-TRIBUNE*, ST. LOUIS POST-DISPATCH, CINCINNATI ENQUIRER, BUFFALO NEWS, DES MOINES REGISTER, *STAR TRIBUNE*, *PLAIN DEALER*, *DAYTON DAILY NEWS*, *NEWSDAY*, *MANHATTAN MERCURY*, *COLUMBUS DISPATCH*, *COLUMBIA DAILY TRIBUNE*, *WILLOUGHBY NEWS-HERALD*, *BELLEVILLE NEWS-DEMOCRAT*, *MANAGED HEALTH CARE NEWS*, CBS-KWTV, WAYNE STATE UNIVERSITY AND CREIGHTON UNIVERSITY.

A SECOND SEMINAR IN KANSAS CITY IS TENTATIVELY SCHEDULED FOR SEPTEMBER 1998. CHECK THE IRE WEB SITE FOR THE LATEST ON UPCOMING WORKSHOPS AND SEMINARS AT WWW.IRE.ORG

Evaluating educators

JIM LUDWICK ATTENDED NICAR'S BASIC BOOT CAMP IN JANUARY AND IMMEDIATELY BEGAN APPLYING THE SKILLS HE LEARNED. IF YOU HAVE RECENTLY STARTED USING CAR TECHNIQUES, LET US HEAR ABOUT YOUR EXPERIENCES. E-MAIL BRENT JOHNSON AT BJOHNSON@NICAR.ORG

By Jim Ludwick
The Missoulian

My first project after boot camp began on the day I returned to the office. Rod Murphy, the associate editor, told me that our education reporter had been trying to determine how various schools had compared in their standardized achievement-test results, that had just become public under a new state law.

The reporter, Gary Jahrig, had been overwhelmed by paperwork he'd received from the state Office of Public Instruction, nearly 200 pages, covering many of the schools in our circulation area, and wasn't sure how to make sense out of the data. I suggested that we could get the Office of Public Instruction to give us the file for the entire state of Montana in computerized form.

The package was published on Jan. 25. My last day at boot camp was Jan. 9.

Training results

I'd attended training at NICAR in early January. I had previous experience, but the NICAR training had been useful for several reasons: it helped me learn to use the latest software; it gave me experience in combining databases; and it provided valuable advice about obtaining and using data from government sources.

When I contacted the Office of Public Instruction, I negotiated an appropriate price for the material: no charge.

The data was on a spreadsheet file, for Lotus 1-2-3. I moved the material to Excel, and from there to Access.

I wanted to identify the schools from particular counties, but the Office of Public Instruction had used code numbers instead of county names. I got a list of the code numbers for each county, and quickly created a database with three fields: the code number, the county name and a designation that indicated whether the county was in the circulation area of our newspaper.

Producing in Access

Once I linked my database to the material supplied, I could easily identify schools in our circulation area and compare their results to the state as a whole. The data covered test results from the fourth, eighth and eleventh grades, showing an average score for each school in each of five subject areas: reading, language arts, math, science and social studies.

Using the Access database manager, I produced tables for each subject at each grade level. My tables listed every school in our circulation area, ranked in order of scores from highest to lowest. The tables included four fields: the school, its county, the score, and the number of students who had taken the test.

Package delivery

A package was published in the Sunday paper. Jahrig wrote the wrap-up story. After interviews with the state superintendent of public instruction and others, he wrote about the test results, the reaction and controversy relating to the use of standardized testing. Another reporter, John Stromnes, wrote a sidebar focusing on one elementary school that had been especially successful. I helped deal with the table and wrote explanatory articles to accompany them. I also made numerous checks and re-checks to make sure I had not asked Access the wrong questions.

The package was published on Jan. 25. My last day at boot camp was Jan. 9.

I already have three other database projects underway. Two involve data I brought home from NICAR.

Jim Ludwick can be reached at
(406) 523-5253 or by
e-mail at jimludwick@marsweb.com

ON THE INTERNET

Soup's On

By Nora Paul
The Poynter Institute

To keep up with this month's food inspection stories, this column will link you to good sites with background, glossaries, and resources about food inspections, food-borne illnesses and other things that will help you cover these important topics and help those of us who made dieting resolutions for New Year's very glad to never eat anything ever again!

The Bad Bug Book
http://vm.cfsan.fda.gov/~mow/intro.html

The handbook on food-borne pathogenic microorganisms and natural toxins.

Iowa State University Extension: Food Safety Project
http://www.exnet.iastate.edu/Pages/families/fs/homepage.html

This has got it all. Read *Food Safety News* daily for story tips and leads and use their list of links to other food safety news sources. Learn how to join the Food Safety Net listserv for daily updates. Use their excellent compilation of other food safety sites to jump to for background and explainers on different food safety issues.

U.S. Food and Drug Administration: Center for Food Safety and Applied Nutrition
http://vm.cfsan.fda.gov/list.html

A well organized collection of pamphlet type material on all aspects of food safety and handling. Good consumer background and tips as well as technical explanations of some food-borne illnesses and concerns.

USDA/FDA Food-borne Illness Education Information Center: Food Safety Index
http://www.nal.usda.gov/fnic/foodborne/fbindex/index.htm

Get the full-text of the FDA 1997 Food Code and a nice list of listservs to join.

Glossary of Epidemiologic Terms
http://www.cvm.uiuc.edu/courses/vp350/glossary.html

If terms like "Extrinsic incubation period" and "gold standard" and "horizontal transmission" are being thrown at you during the interview, be sure you know what they are referring to by using this glossary.

Check out a couple of the **Poynter Institute's** Hot News / Hot Research items on food-borne illnesses and outbreaks, researched by David Shedden:

- Hong Kong Bird Virus
http://www.poynter.org/research/hr/hr_010598.htm
- Tainted Strawberries Served in School Lunch Lead to Hepatitis Outbreak
http://www.poynter.org/research/hr/hr_040397.htm
- Mad Cow Disease
http://www.poynter.org/research/hr/hr_032696.htm

Nora Paul can be reached at
(313) 821-9494 or by e-mail at
npaul@poynter.org

IF YOU WOULD LIKE TO SEE TIPS ON HOW TO BEST UTILIZE THE INTERNET FOR A PARTICULAR TOPIC, SEND AN E-MAIL TO NORA PAUL AT NPAUL@POYNTER.ORG OR BRENT JOHNSON AT BJOHNSON@NICAR.ORG. ALSO VISIT THE POYNTER INSTITUTE'S WEB SITE AT WWW.POYNTER.ORG FOR INTERNET TIPSHEETS.

IRE/NICAR NOTES

Upcoming Events

March 23 - 24, 1998 -- Washington, DC
NICAR/Medill Intermediate CAR Workshop
March 27-29, 1998 -- New York, NY
IRE Broadcast Workshop
May 10-15, 1998 -- Chapel Hill, NC
NICAR Advanced Boot Camp
May 17-22, 1998 -- Columbia, MO
NICAR Basic Boot Camp
June 4-7, 1998 -- New Orleans, LA
IRE National Conference featuring panels for every beat, TV Show and Tell and optional day of CAR
June 14 - 16, 1998 -- Washington, DC
NICAR/Medill Basic CAR Workshop
July 12-17, 1998 -- Columbia, MO
NICAR Basic Boot Camp
July 24-27, 1998 -- Miami, FL
On-the-road training for the National Association of Hispanic Journalists Convention '98
July 29-August 3, 1998 -- Washington, D.C.
On-the-road training for the National Association of Black Journalists Convention '98
August 5-8, 1998 -- Chicago, IL
On-the-road training for the Asian American Journalists Association Convention '98
August 9-14, 1998 -- Columbia, MO
NICAR Basic Boot Camp
December 10 - 12, 1998 -- Washington, DC
NICAR/Medill Basic CAR Workshop

IF YOU HAVE ANY QUESTIONS ABOUT AN UPCOMING IRE OR NICAR EVENT, OR WOULD LIKE TO REGISTER, PLEASE CONTACT OUR MAIN OFFICE AT (573) 882-0684 OR E-MAIL INFO@NICAR.ORG.

Wipe away math fears

SARAH COHEN'S STATS COLUMN THIS MONTH IS A CONTINUATION OF HER ADVICE ON MATH SKILLS NEEDED BY ALL REPORTERS. IF YOU HAVE TIPS OR SUGGESTIONS ABOUT WHAT NUMBERS REPORTERS OUGHT TO KNOW, OR HOW TO OVERCOME MATH PHOBIA, CALL HER OR SEND HER AN E-MAIL.

By Sarah Cohen
NICAR

You probably fear math and numbers far less than you suspect. Most reporters can figure a tip on a restaurant tab. They can split the bill easily. Many invest, buy homes, and work with other complex financial calculations regularly. And, unlike me, most can and do balance their checkbooks.

Here are some tricks to get you started in your quest to conquer your math phobia.

Trick #1: Round off.

If you're talking about the population of the U.S., don't think of 270 million. Think of 250. Or 300. That makes it easy to get a gut feeling of scale. If 12 percent of the U.S. is African-American, round that off to 10 percent, or 15 percent. Now just figure a tip on a $30 dinner bill – somewhere between $3.00 and $4.50 if you're frugal. So the number you're seeking is somewhere between 30 and 45 million people.

The lesson: Think in small, round, numbers that correspond to something you'd be able to figure out in real life: a tip, a 25-percent-off sale, a 10-percent raise. You can change the scale of virtually every number you'll see to get to that point.

Trick #2: Develop a sense of scale.

Another trick is borrowed from John Allen Paulos' book, *Innumeracy*. The book is filled with what Paulos calls simple exercises, which few of us think are simple. It does, though, include one remarkably effective piece of advice: Learn to guess the numbers you look up before you look them up. Being wrong, as you almost always will be, will help cement in your mind the correct answer.

If you guessed that the U.S. had an economy of $450 billion, and you learned through research that it is closer to $8 trillion, you'll remember your error every time you need to use the number. It also means you'll no longer think of a billion as the biggest number you can imagine, the way an editor did when she changed the word 'billion' to 'million' in a story because she thought it was simply too big to be true.

I once admitted to my boss, jokingly, that I didn't pay attention to numbers – they all looked the same to me. He exploded. He rightly mentioned, not so politely, that I couldn't possibly know whether something was right or reasonable if I had no frame of reference. If I didn't know that the Nixon administration slapped on price controls when inflation was only 6 percent a year, how could I possibly call any inflation rate 'small' or 'large'?

You'll have different scales that are important to your beat: How long the typical person lives, the number of people in your community, or the market share of General Motors. Developing your instinct of what's big and what's little will cut seemingly complex issues down to size, eliminating the meaningless numbers and focusing on the important ones.

Trick #3: Watch television news.

Broadcasters have far less freedom to bore readers with meaningless numbers. In any story, you may quibble with the number the reporter chose to focus on. But you'll become more selective in your use of numbers and avoid the all-too-common story construction of anecdote, nut graf, then numbers you hope your readers will avoid.

Trick #4: Ask your best sources for help.

People love sharing certain kinds of secrets. They usually involve secrets that will show you how much they know or how clever they are.

You might ask your county budget director to show you the math (and the spreadsheets) she uses when she develops the budget each year. You might ask the academic author of a new study to walk through some of the methodology with you. You'll become more adept at listening to people talk about numbers – something that has to come while you try to overcome your fear.

Sarah Cohen can be reached at
(301) 942-2199 or by e-mail at
sarah@nicar.org

TECH TIP

Converting dates in Access

By Richard Mullins
NICAR/The Missouri School of Journalism

One of the most common data cleanup needs is converting date information. The two-step process and the concepts for doing this are the same for FoxPro and Access. Arrange the numerals in *mmddyy* order and put in legal separators (hyphens or slashes) between month, day and year. This result may look like a date, but it's still a string of characters that you can't use for calendar math or date arithmetic. Step two is required to convert this correctly arranged date string into a date data type.

The only differences are in the details, which are minor, like the British and the American ways of spelling "realisation." Other Tech Tips have explained the FoxPro steps. This one will show the Access how-to.

The word string is used a lot when talking about computer databases. But it's not a technical term; it's a good metaphor that helps you think about what you can do with strings of characters — cut them apart and tie them together. String functions are what you use for the cutting apart. The ampersand symbol is how you tie pieces together.

You only need two string functions: one that picks out a certain number of characters starting from the left, and another that picks out a certain number of characters starting anywhere in the string.

Assume a table with date of birth information, stored in a character field. I'll call that field **cDOB**, for "character date of birth." The new field to hold the date data-type I'll call **DOB**.

If the **cDOB** column has data like `112184` then step one involves arranging it into `11/21/84`.

Here are the string functions to do that.

```
Left(cdob, 2)    & "/" &
Mid (cdob, 3, 2) & "/" &
Mid (cdob, 5, 2)
```

If the date information looked like `841121`, you'd rearrange the pieces accordingly to get the same result of `mm/dd/yy`.

Step two is converting this character representation of a date into a date datatype.

In Access, this means wrapping CVDate() around the entire command above. Since there are already lots of parentheses, I spread out the two parentheses belonging to CVDate()to make things clearer. (In Access SQL, anywhere you can put one space, you can put several, or even a new line.)

```
CVDate (
Left( cdob, 2)    & "/" &
Mid ( cdob, 3, 2) & "/" &
Mid ( cdob, 5, 2))
```

Putting this fragment together in an update query for our example table (called **Employees**) looks like this:

```
UPDATE employees
SET dob = CVDate (
Left( cdob, 2)    & "/" &
Mid ( cdob, 3, 2) & "/" &
Mid ( cdob, 5, 2) )
```

This works, but we all know that data isn't perfect. Not all combinations of numbers that may have been typed into a government database to represent a real calendar date actually will. We know, and Access knows, that there is no such date as the 30th of February or the 41st of November. If your update query finds any invalid character dates like this, you'll get this message:

```
Microsoft Access can't update all
the records in the update query.

Microsoft Access didn't update ___
field(s) due to a type conversion
failure, ...
```

Click the Yes button to go ahead with the update. All the character dates that couldn't be converted to a calendar date will be set to empty or `Null`. Run a query to show all the rows where the date column you just updated is `Null`.

```
SELECT *
FROM employees
WHERE dob is Null
```

Richard Mullins can be reached by e-mail at richard@nicar.org

Have a suggestion for a Tech Tip or have a technical problem that is stumping you?
E-mail Richard Mullins at richard@nicar.org with your ideas or questions.

Highlights from the Database Library

THE PRICE LISTING FOR ALL DATABASES AVAILABLE FROM NICAR CAN BE ACCESSED ONLINE AT WWW.NICAR.ORG/DATA

IF YOU WOULD LIKE TO SEE A PARTICULAR DATABASE IN NICAR'S LIBRARY DISCUSSED IN DETAIL, SEND AN E-MAIL TO JUSTIN MAYO AT JUSTIN@NICAR.ORG

By Justin Mayo
NICAR

Computer-assisted reporters unite! Now available is a database that NICAR recently updated and that the federal agency overseeing worker well-being maintains.

The Occupational Safety and Health Administration database is cumulative from 1974 through November 1997. NICAR will slice off information from your state or you can order the entire U.S. database on 4 CDs.

The OSHA database includes 4 tables that may be analyzed separately or joined using the activity number field appearing in each one. Here are the tables and some of the information they contain:

• *OSHA:* The main table which gives company sites, previous violations, the number of workers employed at the site in the previous 12 months and other OSHA activity at the sites, such as inspections or complaints.

• *Accidents:* Names those injured, identifies the task that was being performed when the accident occurred, indicates whether a hazardous substance contributed to the accident, lists which body parts were injured, and indicates the degree of the injury – for example, whether it was fatal.

• *Hazardous Substances:* Describes accidents that involve hazardous substances but that do not involve people.

• *Violations:* Details penalties, types of violations – willful, repeat, serious, etc. – and the number of employees exposed to the danger.

Although this relational database structure can be tricky, it should not scare off CAR neophytes. *St. Louis Post-Dispatch* reporter, Virginia Baldwin Hick, first used OSHA data in a *St. Louis Post-Dispatch* series that examined local workplace falls that resulted in fatalities.

"I cut my teeth in computer-assisted reporting by using OSHA data," Hick says. "OSHA had just put into place new regulations on fall protection. I featured how these regulations were supposed to save lives."

Here are some tips Hick, former *Post-Dispatch* reporter Natalya Shulyakovskaya and reporter Russell Carollo, who used OSHA at the *Dayton Daily News*, offer for handling the data.

• Use ZIP codes. "Some companies had many, many locations," Carollo says. "If you used company names, like General Motors, you'd end up with a zillion hits. We used ZIP codes to identify locations."

• Look for deaths. "One of the more interesting queries is to look for citations that result from fatalities," Carollo says. "Once you've got those, you can get all the paperwork on it because somebody's dead. Then rank those in order of monetary fines."

• Find the survivors. "Look through the data field of fatalities, find the names and look up the survivors early on," Hick says. "It helped me focus my story and focus my interviews. We're not talking about meaningless data; we're talking about lives. If you zero in on that story, you see how other things fit."

• Examine penalty fields carefully. "There are actually three different amounts of penalties that you will see for each inspection," Shulyakovskaya says. " First comes from the OSHA table; it's called 'total current penalties' and represents the initially assessed fine. Second is the 'current penalty reflecting changes and modifications', which indicates the fine negotiated between the company and OSHA. It appears in the violations table. Third is the 'total penalty dollars remitted for the inspection,' which what was actually paid. This appears in the OSHA table.

• Take advantage of SIC codes. "We didn't get into looking at OSHA for dangerous occupations because the occupation field is rarely filled in," Hick says, "but standard industrial codes (SIC) are listed. These are standard codes used throughout federal government. SIC codes indicate how the industry is classified in the federal bureaucracy. You can find out which types of industries have the most fatalities."

Justin Mayo is Database Administrator at NICAR. He can be reached at (573) 884-7711 or by e-mail at justin@nicar.org

Health care hints

By Sarah Cohen
NICAR

Whether you care about doctors, managed care companies, insurance or employee benefits, there's no shortage of data sources for stories in the beat. Each piece of the industry is either heavily regulated or heavily supported by taxpayers, making the records readily available. Here's a taste of what's waiting for you.

Hospitals and nursing homes

It's easiest to start with the biggest, most subsidized and often still regulated industries.

Hospital Discharge Databases: The motherlode of data for health care reporters is often in what's known as the "discharge database" for each state. Not all states release it. An alternative is the federal Medicare database of hospital discharges, which includes less detail about doctors but adds an important element – whether the patient was readmitted or died soon after release. Discharge data has been used to document the shift toward shorter hospital stays, to find doctors who would talk about reasons for leaving certain hospitals, to identify list prices, or to figure out which procedures managed care companies will cover.

Hospital financial records: Don't think these are only for business stories. They include important information on the staffing levels of hospitals, the amount of charity care given, and you can estimate how much a nonprofit actually spends – rather than charges – on the care it gives away in return for its special status. Often available through your state health agency or a hospital association in a simpler form.

Medicare cost reports: Any Medicare service that's based on costs rather than fixed prices has become a windfall for some companies. They get paid on the honor system. Then they submit the costs to justify it. Medicare sells these records.

Medicare labor rates: Average wages in each area are used to determine how much hospitals in each area are paid under Medicare. You can use them to figure out how much (or little) employees in your area are making in hospitals.

Insurance records

These are all under state law. You'll often have to type in the simpler records, like financial reports, yourself. But others are available on tape or database in many states.

Malpractice payouts: This is a list of every malpractice claim paid through insurers. In Tampa, reporters found a doctor who had been sued repeatedly over a controversial orthopedic procedure, which included washing bones he used in the operation with a garden hose in his backyard.

Worker's compensation claims: Some private vendors sell acess to workers' comp claims by employer.

Financial records: Any Medicare HMO and many Medicaid HMOs are required to maintain a certain number of private-pay patients, usually through employers. As one official put it, "private industry will demand standards" that regulators can't enforce. Look for HMOs that depend too heavily on Medicare payments. If nothing else, Medicare can cut them off from the program.

Physician and other professional records

Yes, the National Practitioners Databank is closed to us. But some companies are compiling records held in it and selling them to anyone who wants to check on disciplinary actions against their doctors in other states. Other databases that may come in handy are:

UPIN numbers: These are the key identifiers that doctors get so they can get paid by Medicare. Having UPINs will open a world of information to you, especially if you get billing records. The tape costs about $200 from the Health Care Finance Agency.

Corporate records and partnerships: As physicians form associations to vie for managed care contracts, you can learn more about their business dealings by looking through your state's corporate records. Check the 990 summaries at www.irs.ustreas.gov and look under Statistics of Income for the Exempt Organization master list.

Sarah Cohen can be reached at
(301) 942-2199 or by e-mail at
sarah@nicar.org

THIS IS AN EXCERPT FROM THE IRE NATIONAL CONFERENCE IN PHOENIX IN JUNE, 1997. TO ACCESS THE PHOENIX HANDOUTS, VISIT IRE'S WEB SITE AT WWW.IRE.ORG/RESOURCES/CONFERENCES

TO ORDER A COPY OF ANY HANDOUT THAT APPEARS IN *UPLINK*, CALL THE IRE RESOURCE CENTER AT (573) 882-3364.

CHECK OUT MEDICARE'S WEB SITE AT WWW.HCFA.GOV, THEN CHOOSE THE DATA AREA TO DOWNLOAD A LIST OF DOZENS OF DATABASES MEDICARE SELLS TO HEALTH CARE CONSULTANTS.

Bits, Bytes and Barks

Membership Message

You can join IRE via our Web server, or you can download a copy of our membership form in Adobe Acrobat format from the Web site and submit it by fax to (573) 882-5431.

For more information, check out http://www.ire.org/membership.html

IRE Memberships are $40 for professional journalists and journalism educators, $25 for students and $55 for international journalists. For a detailed account of what IRE has to offer, contact the IRE main office at (573) 882-2042.

NICAR Net

Talk in never cheap on the IRE and NICAR listservs. A wealth of topical discussion awaits you.

To subscribe to IRE-L or NICAR-l, send e-mail to listproc@lists.missouri.edu

In the body of the message, type:

subscribe NICAR-L<your name>

subscribe IRE-L<your name>

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

Moving on

IRE board member Shawn McIntosh has moved from *The Dallas Morning News* to become managing editor at *The Clarion-Ledger* in Jackson, Miss. If you have recently switched locales, let us know of your whereabouts.

Campaign Finance

Jack Dolan, director of the Campaign Finance Information Center, is continuing to update the CFIC Web site at http://www.campaignfinance.org.

Campaign contribution data is now available from 6 states – Illinois, Indiana, Kansas, Minnesota, Ohio and Wisconsin. The data can be easily downloaded, and comparisons can be made by joining tables between the state files. CFIC hopes to enable such comparisons to be made online once common fields are established and standardized. CFIC's overriding goal is to acquire campaign contribution data from as many states as possible to allow national contribution patterns to be discerned.

The CFIC Web site also has added new stories to its story library, links to related Web sites and the inaugural issue of the newsletter Tracker.

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

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