Uplink

September • October 2004
Volume 16, Number 5

Published bimonthly by the National Institute for Computer-Assisted Reporting

www.nicar.org

ELECTIONS

Purged voter problems

By Chris Davis and Matthew Doig
Sarasota Herald-Tribune

Our analysis of a controversial database that Florida election officials were using to purge felons from the voter rolls exposed a fatal flaw that led to the list's demise. The analysis was simple, though the story was anything but ordinary.

We imported the state's data into Microsoft Excel. With the help of a pivot table, we found how many of the 48,000 people on the list were white, black or Hispanic. Excel's pivot table function allows users to create cross-tabs to categorize and summarize data.

The check was a basic step any journalist should take after he or she first gets new data.

Immediately, the problem appeared in the results – 24,000 of the potential felons were white, 22,000 were black, and 61 were Hispanic. We knew that Hispanics in Florida make up a key Republican voting

continued on page 12

SPOTLIGHT: EDITING & MANAGING

Guiding the project from start to finish

By Tim Ragones, *IRE and NICAR*

Working on a computer-assisted reporting project can mean dealing with slow record requests, dirty data and uncooperative software. But untrained editors may pose as much an obstacle to projects as these other hurdles.

That's what John Ullmann, IRE's first full-time executive director and the current executive director of the World Press Institute in St. Paul, Minn., told a room of newsroom managers at a CAR boot camp for editors earlier this year in Columbia, Mo. If project reporting is a team sport, Ullmann said, editors are its weakest members.

Ullmann, tapping into his experience as a former project editor, provided guidance to the editors and news managers at the boot camp on how to build better computer-assisted stories from conception to publication. Ullmann served as the assistant managing editor for projects at the *Star Tribune* in Minneapolis for six years. He supervised projects that won more than four dozen awards. In 1990, a series showing how members of the St. Paul, Minn., fire department profited from arson won the Pulitzer Prize for investigative reporting.

In Ullmann's experience, the typical computer-assisted or investigative project lasts three to six months.

continued on page 22

SPOTLIGHT:

For more about managing and editing see:

- Why it's important for news managers to understand CAR, p. 2
- Resources for editing and managing CAR stories, p. 22
- The September-October *IRE Journal*

Starting small

By David Donald, *IRE and NICAR*

What's next?

That's the question I face most often when I finish teaching an introductory session of computer-assisted reporting to journalists from small news organizations.

What they're often really asking is simple: How do I go back to a small- or medium-sized newsroom that has never done CAR?

Sad to say, even today some larger newsrooms exist without doing much CAR. But for reporters in smaller orga-

continued on page 20

Bits & Bytes

Hazardous-site data available

Journalists investigating environmental issues in their community can use a database about hazardous waste sites now available in the IRE and NICAR Database Library. The database contains information on potential and confirmed hazardous waste sites at which the U.S. Environmental Protection Agency Superfund Program has some involvement.

Maintained by the EPA, the Comprehensive Environmental Response, Compensation and Liability Information System (CERCLIS) database has been cleaned by NICAR data analysts for easier use. It includes two tables about contaminants and responsible parties that are not available for download from the EPA Web site. The data is current as of April 2004.

To order the CERCLIS database, contact IRE and NICAR at 573-884-7711.

Pollution beat book

IRE recently published an environmental pollution beat book, "Covering Pollution," in collaboration with the Society of Environmental Journalists. The book shows journalists how to get to the heart of an investigation quickly and without waiting months for Freedom of Information requests to be fulfilled. Included is information about how to use Web-based, federal database searches and tips from veteran reporters about how to handle investigative reporting aspects such as interviews and writing. To buy the book or to search for another IRE beat

continued on page 4

INSIDE NICAR

Putting a focus on the editing

By David Herzog, *NICAR and Missouri School of Journalism*

Here at NICAR and in the pages of *Uplink*, we tend to focus on computer-assisted reporting. It makes sense, given that the majority of journalists who are doing CAR are reporters.

But reporting is only one part of the process, and the reporting side is no more important than the editing. So with this issue of *Uplink* we provide guidance for editing CAR stories and address some of the unique challenges these stories present.

Also, you'll find articles about how to start CAR in smaller newsrooms – something that's on the mind of many journalists in these days of fewer resources.

Editors and news managers have specific needs for hands-on CAR training. That's why IRE and NICAR regularly offer a mini-Boot Camp for editors. Ten editors from newspaper, television and radio attended the last session. They learned how to make CAR successful in their newsrooms, negotiate for electronic records, and understand what their reporters are doing.

Next year's Editor Boot Camp is April 8-10 in Columbia, Mo. You'll find registration materials at *www.ire.org/training/april05EditorCamp.html*. If you're an editor or news manager who wants to understand more about CAR and allow it to blossom, this is essential training.

Contact David Herzog by e-mail at dherzog@nicar.org.

Upcoming hands-on CAR training

Looking for intensive computer-assisted reporting training in the coming year? IRE and NICAR have scheduled several introductory and specialty hands-on CAR seminars.

There are four weeklong Boot Camps in Columbia, Mo., for journalists who want to learn how to acquire electronic information, use spreadsheets and databases to analyze the information, and to translate that information into high-impact stories. IRE and NICAR provide follow-up help after participants return to their news organizations. The 2005 Boot Camps are Jan. 9-14, March 20-25, May 15-20 and Aug. 7-12.

Journalists interested in learning how to map data for news stories can take advantage of mini-Boot Camps in using geographic information systems (GIS). The training sessions, Jan. 14-16 and Aug. 19-21 in Columbia, will use ArcView GIS. Participants will learn how to create basic maps, work with projections, and geocode data. Additional GIS training is being planned for Washington, D.C.

Hands-on classes will be offered at the Annual CAR Conference March 17-20 in Hollywood, Calif., and at the IRE Annual Conference June 2-5 in Denver.

For complete information about IRE and NICAR training visit *www.ire.org/training*.

INSIDE NICAR

Scraping away for useful data

By Derek Willis, *The Center for Public Integrity* and Jeff Porter, *IRE and NICAR*

Been scraping lately?

Toss out the mental images of home improvement, and instead think about databases on the Internet. The rise of government Web sites provides a great opportunity for CAR-savvy newsrooms. As with many developments involving governments, the promise is somewhat hidden, but IRE and NICAR are planning a project to change at least part of that.

More and more public databases are being placed on the Web, but often they are hidden behind a search form, or they require hours of repetition to save each record. And then there's the time spent separating the data from the HTML that surrounds it.

The Scraping Project is an answer to these problems. With a small dose of programming (think scripting languages like Perl, PHP or Python), news organizations can "scrape" the information directly from a Web site and save it in a format that spreadsheet and database programs can use.

Many state and federal agencies now provide data in this way, even if the underlying data can be obtained through public record requests. The advantage of scraping is that often you can get the data in a much quicker time frame.

Here are some recent examples of scraping: Mark Schaver of *The* (Louisville) *Courier-Journal* wrote a script in Perl to grab information about state-registered charities from the Kentucky attorney general's Web site. The script loads more than 2,600 pages and pulls out the information to create a tab-delimited text file with the various fields displayed. You can see the script at *www.open-sourcecar.org*, then click on the link near the top about the "First Scrape."

In another instance, John Perry of *The Oklahoman* used Perl to scrape docket information off a state court Web site for data that helped form the basis of a story detailing a county judge's habit of exonerating forfeited bonds. (For details about Perry's scraping efforts, see the Tech Tip on page 16). In both cases these scripts performed the sort of drudge work that we often don't have time for or can't do without the risk of making a mistake (or several). This is what computers are made to do.

Other IRE members have done – or will soon do – similar tasks, and the Scraping Project will become a central resource for either the code or the data (or both, depending on the situation). Journalists would identify data sources in their states or cities that are not available as a single download, and then either write a script themselves or ask for assistance from fellow scrapers.

Journalists can make their own copies of certain public databases and check to see if the files have been updated or the search form has changed. Participants can use any tools or languages they want to accomplish the task; the goal is to get the example posted and available for use or conversion for different scraping scripts.

Depending on the nature of the data and competitive factors, the script – and perhaps the data – would be made available to other journalists through IRE's Web site.

For news about the Scraping Project, keep checking the main IRE Web page (*www.ire.org*). Developments will also be announced on the NICAR-L listserv.

There are, of course, certain barriers to overcome in a project like this. The first is that as journalists, we're competitive people who don't always like to share the fruits of our labor. But the CAR community has a tradition of helping each other and even providing code or data to others in need, and this idea builds on both.

Now, to spark some conversation among the CAR community about this idea: What appeals to you (or doesn't)? How would you improve the project? IRE is uniquely positioned to make such an effort a success, making more data and skills available to a wider range of reporters and editors.

Contact Derek Willis by e-mail at dwillis@gmail.com.

Contact Jeff Porter by e-mail at jeff@nicar.org.

MOVING?

Please send us your new address so you don't miss one issue of *Uplink*.

Drop us a note at
jgreen@ire.org

Investigative Reporters and Editors, Inc.

Bits & Bytes

continued from page 2

book, visit *www.ire.org/store/books* or call 573-882-3364.

Hazmat updated

New data in IRE and NICAR's Hazardous Materials Incident Reporting System (HMIRS) database can help journalists investigate the impact of accidents in their region. The database includes incident reports about hazardous material releases for all types of transportation (air, highway, railway and water).

The main table, Hazmat, provides incident details including location, carrier, shipper and consigner's identity, origin, and destination. Data covers the period between 1970 and June 3, 2004.

The IRE Resource Center offers examples of hazardous material stories including one by the *The* (Cleveland) *Plain Dealer* that found the Federal Aviation Administration had failed to monitor hazardous materials transported on airplanes (Story No. 15476). For sample stories, visit the IRE Resource Center at *www.ire.org/resourcecenter*. For more information or to order the HMIRS database, contact IRE and NICAR at 573-884-7711.

Editor's note

In the last issue of *Uplink*, "Finding faults in new homes" by Katy Miller of the *Orlando Sentinel* included an incorrect name. It should have referenced senior projects reporter Dan Tracy. Construction stories in which Tracy took part can be found in the IRE Resource Center (*www.ire.org/resourcecenter*) under Story No. 21240 and 20682.

CRIME

Cobbling together data for public lands

By Judd Slivka, *The Arizona Republic*

Our story about crime in parks along the Arizona-Mexico border started after a park ranger remarked to me, "It's getting so dangerous out here that I won't even take my family camping anymore."

At first, my newspaper wanted to measure crime on public land in Arizona and the West, but we quickly ran into obstacles. One big problem: Five federal agencies dominate land ownership in the Southwest, and none kept their data the same way. The reporting got complex, even for what seemed like simple questions. What do we measure? What is "public land?"

For an agency like the National Park Service or the Fish and Wildlife Service, there's an easy answer: Public land is land that the public pays for and gets to use. But then we hit the Bureau of Land Management, a huge player in the Southwest controlling 14 million acres in Arizona alone. Most of that land, it turns out, is grazing land owned by the public but not open to the public. Working with BLM's public affairs and law enforcement staff, we finally limited the scope to properties in the National Resource Conservation System, which includes national monuments and national conservation areas.

Even then, there was so much public land that we ended up limiting the scope even further by focusing on the region west of the Rocky Mountains.

We had to deal with five agencies, none of which kept its data in the same way. We knew that the easiest way to measure crimes would be to use FBI Uniform Crime Report figures. The federal agencies, however, didn't record their crimes using the UCR. Or if they did, the agencies did not keep them in a format that was easy to export.

The USDA Forest Service, for example, has a new system for reporting crimes that uses a violation code that does not conform with UCR categories. It took days to figure this out by building an Excel spreadsheet for each of the 40 forests in the West and then scanning the UCR handbook to find the corresponding code.

The National Park Service has a standard form that its parks are supposed to use and are designed to work with the UCR codes. The standard form is so good that we adopted it to collect data. Then we discovered that not all parks use the form. For instance, parks in the Intermountain Region, which includes the Southwest, use a variation of the form. The Midwest Region does not use the form at all.

We couldn't tell what the Fish and Wildlife Service uses because the agency took two months to provide data for the 10 refuges in Arizona, and it provided only the barest information.

Making rates

Now that we had the data, we had to figure out how to measure it. The statistical part was simple; we would use a common benchmark, crimes per 10,000 visitors. Though we knew most crimes weren't committed by registered visitors, especially near the U.S.-Mexico border, it gave us a consistent measure. We also chose to measure acres per law enforcement officer, which gave us another way of looking at public lands in the open spaces of the West.

We noticed another problem: the agencies not only indexed their data differently, they responded to crimes differently. A murder at a national park would most likely be investigated by the park service and re-

ported in that agency's statistics. A murder on BLM property, though, would probably be investigated by the local sheriff's office.

After we compiled the data and sorted it in Excel, we faced a new problem: dissimilar data. The numbers looked the same, but they were not comparable. We decided we couldn't make a comparison between agencies. We used the combined list internally to show us where to focus the reporting. For readers, we published "Top 10" lists of units with the most crime for each agency, plus a map of Arizona with the different crime statistics for each property.

With the data in hand, I started spending time with rangers in the field, and we were able to substantiate most of our findings with anecdotes. It took about 10 weeks to complete the story while working on other assignments.

Our story was much more limited in scope than we had originally set out to do, but very compelling. Our database guided us to a story on how crime on the Arizona-Mexico border was drawing law enforcement officers of the agencies away from their assigned stations. As a result, crime was going unnoticed and unpunished on vast tracts of recreational land.

We used the statistics for the West to add perspective and show similar high numbers at northern border parks, such as Olympic National Park in Washington state.

The database wasn't perfect. But it was a great guide for us to write a story on a subject close to the hearts of our readers.

Read "Border crime ravaging parks in Arizona" at *http://www.azcentral.com/specials/special03/articles/1026publiclands26.html*

Contact Judd Slivka by e-mail at judd.slivka@arizonarepublic.com.

CALL FOR ENTRIES

IRE AWARDS 2004

The annual contest of Investigative Reporters and Editors, Inc.

DEADLINE: Postmark by Jan. 10, 2005

The contest recognizes the best investigative reporting in print, broadcast and online media, and helps identify techniques and resources used by entrants.

Categories

Newspaper:
For outstanding investigative reporting at a daily or weekly newspaper. (Use highest one-day circulation of the week.) Categories are: Circulation less than 100,000; Circulation between 100,000 and 250,000; Circulation between 250,000 and 500,000; Circulation more than 500,000; and Local-Circulation Weekly.

Television:
For outstanding investigative reporting by a television outlet. Categories are: Network or syndicated program; Top 20 market; and Below Top 20 market.

Other Media:
For outstanding investigative reporting in other media such as Magazine, Newsletter, Specialty Publication, Book and Radio.

Online:
For outstanding investigative reporting.

Special Categories:
Tom Renner Award: Outstanding reporting in any medium covering and exposing organized crime. (Limited to 10 stories) You must submit TWO sets of copies if submitting an entry in the Tom Renner category and in another category.

IRE FOI Award: Honors individual or organization in any medium whose significant actions further open records or open government. You must submit TWO sets of copies if submitting an entry in the FOI category and in another category.

Student Award: Outstanding investigative reporting by a student in a college-affiliated newspaper, magazine or specialty publication (or while serving a print internship), or broadcast work that has been publicly reviewed, screened or aired.

International Entries: International entries will be placed into appropriate categories by IRE staff. Contest judges can then move entries into other categories. IRE can award a special citation for deserving international work.

NOTE: Judges reserve the right to give more than one award in a category or to declare no winner in a category

For entry forms and additional information, visit our Web site at
www.ire.org/contest

MAPPING IT OUT

The latest uses of mapping in news reporting.

Plotting firehouse coverage

By Thomas Gaumer
The (Cleveland) Plain Dealer

As part of a regionalism series, *Plain Dealer* reporter Martin Stolz decided to explore fire services. A central question was whether greater Cleveland had duplicative fire departments, equipment or stations. He quickly realized the fire stations element was a story best told visually.

Cuyahoga, the county around Cleveland, has a total of 103 fire stations. Stolz wanted to see if the county really needed that many, or if perhaps consolidation might save taxpayers millions of dollars without affecting firefighting abilities.

Stolz discovered that Cuyahoga County exceeds the national median by 47 stations, 42 ladder trucks and 67 pumper trucks. To tell the story in more than just the dull numbers, Stolz turned to mapping.

The *Plain Dealer* was making the switch from ESRI ArcView 3.3 to ArcView 8.3 and this was the first chance to do a major map with the new software. A lot of what we did was trial and error to find out what worked best to get the map to the graphics department.

The result was a map of the location of every fire station in Cuyahoga County with a one-mile radius around it. We decided on a mile, even though fire stations cover far more than a mile, because those circles showed the proximity of stations and where they overlap, particularly in the southeast area of the county. Smaller circles also avoided a cluttered map.

Generating the map was easy.

We wanted ArcView to geocode all the fire stations so we would have a dot on our map for each station. But as anyone who has done geocoding – creating map points based on street addresses stored in a data table – can attest, you rarely get 100 percent results. But we needed 100 percent.

To tell the story in more than just the dull numbers, Stolz turned to mapping.

To geocode, we used a map from the Street Map extension we had purchased with ArcView 3.3. Using 3.3, we made a shapefile of streets in the *Plain Dealer's* coverage area, and ArcView 8.3 read that file. We set up a geocoding service based on that file.

Stolz had created a Microsoft Excel spreadsheet with every station in the county. He called each community as he had time and typed in the address, city and ZIP code. He also assigned each a station identification number. Then it was simply a matter of converting this Excel file to dBASE, which ArcView can open.

On our first pass through geocoding, ArcView plotted all but 15 of the 103 stations. That's when the hard part started.

For several addresses, ArcView had the street name, but not an exact address, and we were able to force ArcView to geocode those addresses using a nearby address. Because of the size of the map, the reader would never be able to tell we had slightly moved a station's address.

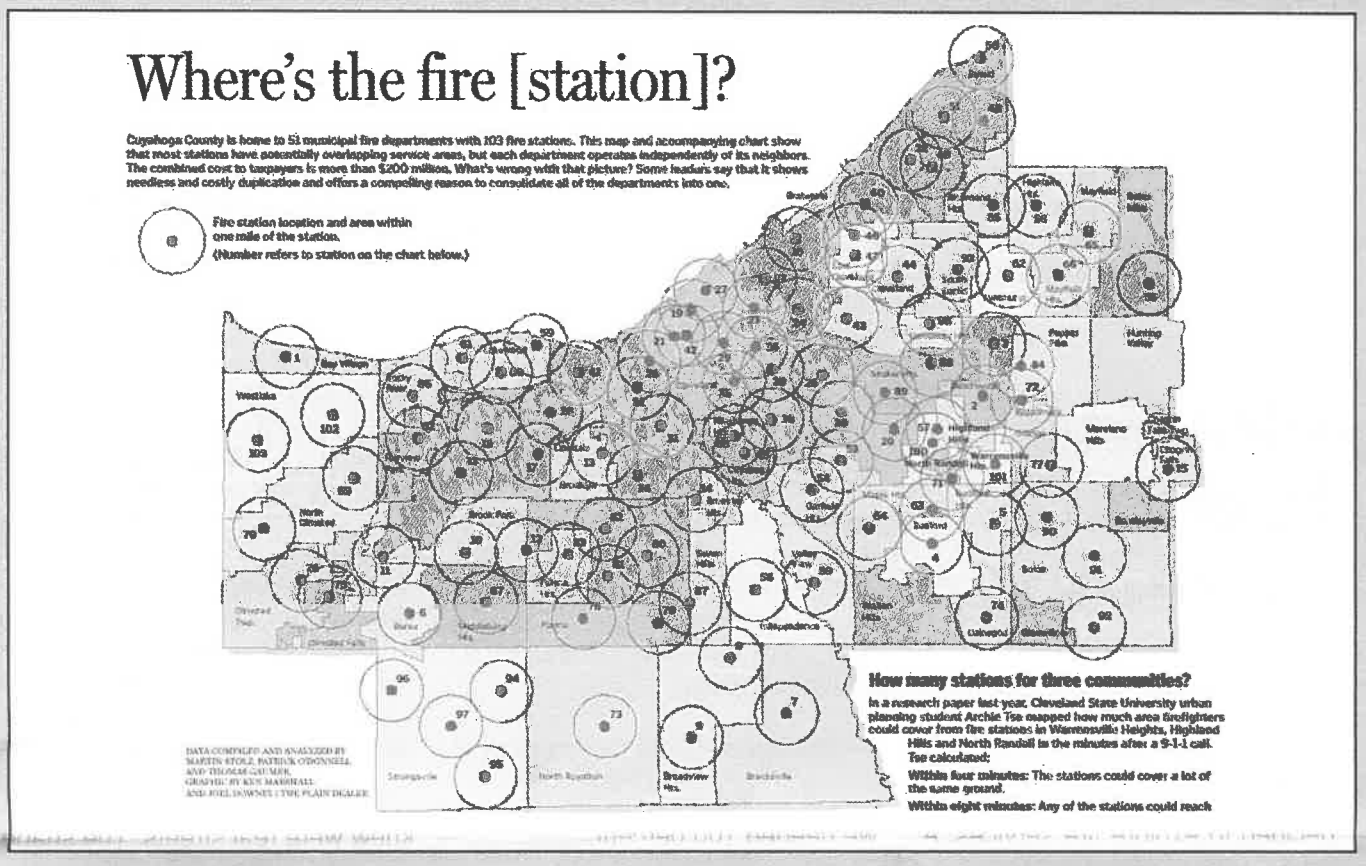
Then it got harder, and there were still a half dozen or so addresses where there was no match at all. ArcView just didn't have any suggestions.

So we turned to the data in the streets shapefile to determine why ArcView would not find what we knew were real streets. The shape file had thousands of records so we used ArcView's sort function to alphabetize the names of the streets.

Sure enough, we found one street that was not even in ArcView. It was barely more than an alley, yet it was the legal address of the fire station. We proofread the map carefully and noticed that ArcView had put one station in the wrong suburb, although it was practically on the city line.

We turned to the paper copy of the reverse directory and looked up a known address on a major street very close to that alley. So we deliberately gave some stations the wrong address from the standpoint of the post office, but perfectly fine for ArcView.

In the case of the station that was in the wrong suburb, we had to assign a new address that was slightly changed to move it into the correct suburb. Because we did not



want the station to appear twice, we used ArcView's edit table function to delete the old record from the database. Once it was deleted, the dot disappeared.

Perhaps the most puzzling to us, however, was ArcView's failure to assign stations or even to see a road named SOM Center. This is a major state route in Cuyahoga County, and it didn't make sense that it did not exist in ArcView.

SOM stands for "Solon-Orange-Mayfield Center." The letters are officially separated by periods, as in "S.O.M.," though we entered it without, as is common practice.

When we turned to the dBASE file with the fire stations, we found that ArcView had changed the name of the road to "S. O M Center Rd." The S should have been part of the name, but to ArcView, it was "South." The "OM" in SOM was seen as "O M," with a space between the letters.

Once we found the problem it was easy enough to fix. We just changed the addresses of the stations that would not geocode to S. O M Center Road. ArcView geocoded them easily.

We had gotten a map of all the communities of Cuyahoga County from Cleveland State University and added that as a layer.

Stolz had numbered all the stations, so we told ArcView to label the stations with the corresponding number on a large chart with information about each fire department.

It was easy to draw the one-mile radius around each station using ArcView's buffer wizard and telling it we wanted a buffer of one mile. At first, ArcView made circles appear as ovals, which signaled that we needed to pick a better map projection. After we selected a projection for northern Ohio the ovals turned into circles.

We exported the map in the Adobe Illustrator format and our graphics staff was able to open it with Illustrator and change colors and reposition numbers and names as they wanted.

Contact Thomas Gaumer by e-mail at tgaumer@plaind.com.

Would you be willing to share a mapping example with fellow journalists? Send an electronic copy of the map along with details to David Herzog at dherzog@nicar.org

FIRST VENTURE

Mapping hidden patterns in crime-ridden city

By Brad Branan, *Tucson Citizen*

I found an investigative story last spring before I even moved to Arizona to work as a projects reporter at the *Tucson Citizen*.

The Tucson metro area had the nation's highest crime rate, Sperling's BestPlaces found in a study of FBI data.

It wasn't the ranking that intrigued me, though. It was the Tucson Police Department's response.

Department officials gave an explanation that will be familiar to many police reporters: They blamed the FBI's Uniform Crime Reports for providing a distorted picture of crime.

The UCRs equate a rape with the theft of a garden hose, a department spokesman said, in what became a mantra for city officials.

The statement was true on its face but also suggested that Tucson didn't have a problem with violent crime, and residents didn't need to worry about safety. That seemed unlikely.

Using Microsoft Excel, I crunched the FBI numbers that I downloaded from the FBI Web site to come up with rankings for violent crime – murder, manslaughter, robbery, rape and aggravated assault. The violent crime rate for the Tucson area was 55th out of 280 metro areas, ahead of places such as Philadelphia and Washington, D.C.

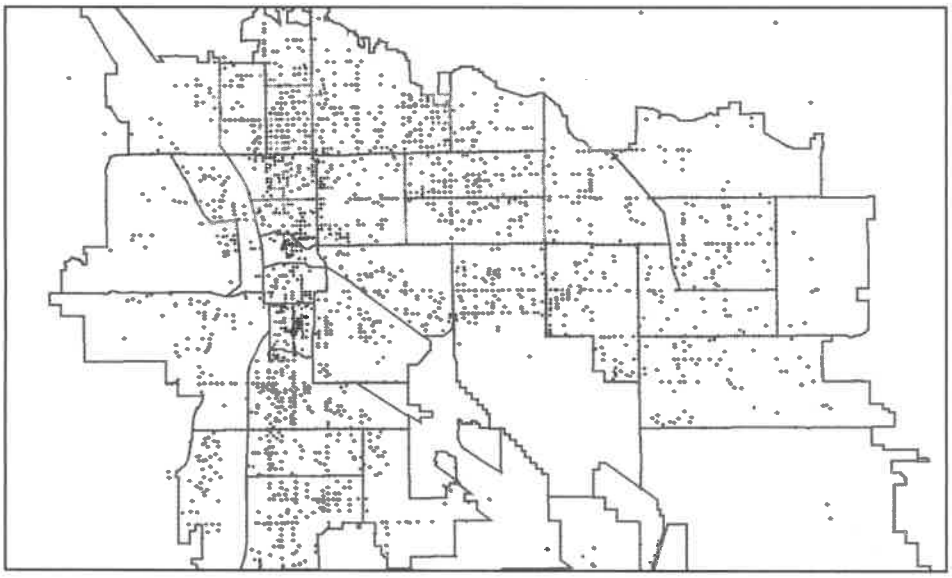
Violence was clearly a problem in Tucson. Still, the numbers didn't reveal the story that needed to be told. I needed to find out where the violent crime was concentrated and what it was doing to the neighborhoods.

The skills I learned at an IRE and NICAR mapping mini-Boot Camp this year helped me answer those questions.

Mapping crime

An electronic map has streets, city boundaries and other elements found in the paper map on the newsroom wall. In a geographic information system (GIS) program, those elements are called layer files.

GIS clearinghouses across the country collect layer files of highways, voting precincts, Census tracts and other important map elements from various government agencies. The files are often available for free on the Internet.



Go to the Center for Advanced Spatial Technology Web site (*www.cast.uark.edu*), click on "Starting the Hunt," and you'll find a list of GIS sites nationwide. That's where I found a county transportation GIS site that had all the map layers I needed. I opened them in my mapping program, ESRI's ArcView 8.3.

I was ready to add the crime reports to the map. Tucson police provided me with an electronic file of the reports but wouldn't give me exact street addresses, saying the information could violate the privacy of crime victims.

I didn't argue because ArcView allows you to map data such as crime reports without exact addresses. The reports were listed by city block, which was specific enough for my needs. For example, a crime that might have happened at 4850 S. Park Ave. would be listed at 4800 S. Park Ave.

I created a spreadsheet of the crime reports in Excel and then opened the file in ArcView. I had the program place a dot on the map for the location of each report, a process called geocoding. The resulting map showed dots throughout Tucson and not a definitive pattern. I had to dig deeper.

To find the areas with the greatest number of crimes, I needed to count the reports within smaller geographic areas. I attached the geocoded crime reports to a map of 38 police patrol areas. ArcView joined the two files, creating a single file of crime reports grouped by patrol area.

I opened the database in Microsoft Excel and counted the number of violent crimes in each patrol area. The results still didn't tell me which area had the most violent crime. Some of the patrol areas covered more area or had larger populations than others.

The crime reports had to be assigned a rate to account for the differences. I based the rate on square mileage instead of people because of rapid and uneven population growth in Tucson since the 2000 Census.

While I was putting together the map, editors and reporters speculated that the highest violent crime rate would be found on Tucson's south side. Drive-by shootings and other forms of gang violence in that part of town receive a lot of headlines and TV news coverage.

The perception proved to be wrong. The area with the highest violent crime rate was on the other side of town – the north side. After confirming my results with the police department's statistics expert, I was ready to report.

Observations

By placing a map of city neighborhood associations over the crime clusters, I knew who to talk to in each community.

The neighborhood activists were defensive at first. They thought my story would detract from recent efforts to reduce crime and redevelop their communities.

They became more cooperative after I explained that I wanted to know about the causes and the effects of the violence, not simply label Tucson's most dangerous neighborhoods.

The north side had long been the site of the city's red-light district, attracting prostitutes and drug dealers who operate around inexpensive motels and strip clubs. The activity has created a volatile mix, producing violence that spills into nearby neighborhoods.

Most victims are involved in some sort of criminal activity, police told me. Wanting drugs or a prostitute, they end up getting robbed, beaten or killed.

Law-abiding residents were also paying a price. The violence convinced many of them to move to safer parts of town, neighborhood activists told me. The activists themselves had been subjected to threats and other forms of intimidation.

Census 2000 provided the data to document their observations. Adding Census tract data to my crime map, I found that only 23 percent of residents in the highest crime areas had lived there more than five years.

The perception proved to be wrong.

The Census also provided some statistical evidence for an explanation police had for the crime clusters: poverty. The areas with the highest crime rates also had some of the city's highest poverty rates.

Feds respond

The ability to downplay violent crime was diminished after the story was published. Police and other city officials acknowledged the destructive effects of violence in the neighborhoods that were profiled.

In June, the Bureau of Alcohol, Tobacco, Firearms and Explosives announced a violent crimes task force for the city. A city official said my story was the reason the task force was formed. An ATF special agent said the story helped build support for the task force.

ATF said it would use "new innovative technology in combination with analytical investigative resources" to reduce crime. An ATF map of "gun-related hot spots" in Tucson mirrored the one published in the newspaper.

The task force promised to bring federal, state and county law-enforcement resources to Tucson, where police department staffing is well below the national average.

Despite a budget crunch, the Tucson City Council approved the police department's request for $4 million to hire 71 more officers. And Attorney General John Ashcroft visited the ATF-led task force here and praised its initial efforts.

Previously reluctant to discuss violent crime, neighborhood activists and city officials said the story was helping to attract needed attention and resources to the area.

"Your report of the prevalent problems in the Oracle-Miracle Mile corridor gives neighbors and businesses hope that people are really paying attention to the needs of the area," a city councilwoman wrote me. "I truly appreciate your work."

Contact Brad Branan by e-mail at bbranan@tucsoncitizen.com.

ELECTIONS

Lobbyist data rich with trips, food, gifts

By Kit Wagar, *The Kansas City Star*

With the kickoff of the 2002 Super Bowl just hours away, a trio of unlikely business honchos found themselves on a corporate jet winging their way toward New Orleans.

It was rarefied air for Wayne Crump, a cattle farmer and former sheriff's deputy from a rural Missouri hamlet; for Crump's wife, Tiffany; and for Jim Foley, a former pipefitter from a working-class suburb of St. Louis.

What put them on a Monsanto Company jet with $400 tickets to the Super Bowl were Crump's and Foley's part-time jobs as majority leader and assistant floor leader of the Missouri House of Representatives.

The entire trip – the private flight, food and drink after the game, and tickets to see the St. Louis Rams lose on a last-second field goal – was worth $4,136.61, as reported by two lobbyists who picked up the tab.

The trip came to light two years later, after *The Kansas City Star* launched a series of articles about lobbyist spending in the Missouri statehouse. The ongoing series is based on the analysis of data about more than 20,000 gifts that lobbyists reported giving to Missouri's elected officials, their staff and family members from January 2001 through July 2003.

The gifts ranged from free meals, concerts and alcohol, to seminars at Harvard, lavish birthday parties, trips to the Kentucky Derby and New Year's at the Four Seasons lodge for seven legislators and their spouses.

The value of the gifts totaled $2.4 million. But what angered many readers was that spending had risen from $738,020 in 2001 to $916,828 in 2003 – the year after term limits had replaced more than half of the legislature.

These were legislators who supposedly were not beholden to special interests, a new breed who would do things differently. The analysis found the major difference was that the new members tried harder to cover up the gifts they received from lobbyists.

But that effort to conceal the actual recipients was a relatively minor difficulty. The tougher problems were the classic ones when working with databases: dirty data, vague or incomplete descriptions, getting lobbyists to answer questions about their reports and the inherent difficulty of undertaking a complicated project while continuing to cover a beat.

Work on the project began in September 2003, when I requested the entire database of lobbyist gifts from the Missouri Ethics Commission. The information technology manager of the commission sent me the data in a Microsoft Excel spreadsheet by e-mail at no charge. The commission started an electronic reporting system in January 2001, so the data covered three legislative sessions.

When I last looked at overall lobbyist gifts in 1999 the records were kept on 3-by-5-inch index cards. I had to photocopy more than 200 index cards then build my own Excel spreadsheet of some 3,400 gifts provided during one six-month legislative session.

The database included gifts received by each legislator and the six statewide officials. It also included gifts to family members and staff members. Before 2001, the Ethics Commission never tracked spending on family members or staff members.

The data, however, reflected the haphazard way lobbyists entered the information.

When lobbyists entered information, they often entered the name in the wrong field. Or they added conflicting information.

The number of ways a single name could be entered was astounding. Michael Gibbons, the Senate majority leader, was entered as Mike Gibbons. Michael Gibbons. Sen. Michael Gibbons. Sen. Mike Gibbons. Senator Gibbons. Sen-with-no-period Gibbons. Gibbons with one "b," Gibbons with three "b"s.

All those had to be standardized because the computer sees each variation as a different person. Any record with a different spelling will be missed by Excel's filter function. And when the records are moved into Access to aggregate the data, false totals and false rankings will result if every entry isn't standardized.

I spent parts of three weeks cleaning up the data and putting it into usable formats. I divided the data into four broad areas: individual elected officials; groups of lawmakers; staff, family and friends; and lobbyists. I created several tables – mostly subsets or compilations of the raw data – so I could analyze it differently. I added an ID number to each official, each staff member and each group. But I didn't need to do any cross-tabulations or other analyses that included joining tables.

Then I added new fields: party affiliation; whether the individual served in the House or the Senate or as a statewide official; the region of the state they represented; and memo fields. The memo fields allowed me to enter information that might affect the analysis, such as when a staff member quit working for one lawmaker and began working for another. It also allowed me to list the titles of executive department employees, whether a gift was solicited by the government official, and the major clients of selected lobbyists.

The process was complicated by term limits, which took effect in 2002. Instead of 197 legislators, I had 294 who served for different time periods.

Some lawmakers had moved from the House to the Senate. For some analyses, such as spending on House members compared with the Senate, I needed to keep those numbers separate. For others, such as an overall ranking of gifts received, I needed them combined.

Once the data was cleaned up, I used Excel to perform calculations on the data and to sort the gifts by size, date and recipient. I exported the information into Access to aggregate the gifts for each official, staff member and family member.

An interesting analysis aggregated the data so it showed how much each lobbyist gave to each official or staff member. For example, the lawmaker who received the most favors among Kansas City's delegation got more than 80 percent of his gifts from a single lobbying firm.

By comparing data from the first seven months of each year, I was able to show trends in spending by party and by committee. After Republicans took control of the House in January 2003, lobbyist spending on the Social Services Appropriation Committee fell 90 percent to $675. Meanwhile, spending on the 14 tax- and business-related panels rose 141 percent to $53,463.

For years lawmakers had formed private groups known as caucuses. When lobbyists reported spending on a group, the final recipients were never reported. In 2003, lawmakers in both parties expanded the use of caucuses as a way to disguise their sponging off of lobbyists.

In 2003, the number of partisan caucuses doubled to 38 from 19. While spending on individual lawmakers fell 34 percent from 2001 to 2003, spending on private caucuses rose 80 percent.

Republicans in particular used this vehicle. In 2001 and 2002, when Democrats controlled the House and Republicans controlled the Senate, lobbyists spent about $103,000 on Democratic groups and $86,500 on Republican groups.

In 2003, after Republicans gained control of both chambers, spending on Republican caucuses skyrocketed to $135,755, nearly three times the amount spent on Democratic groups. Interviews with the caucus chairmen showed that the Republican groups often functioned as private forums for lobbyists to make a pitch.

The total numbers give backbone to a project like this. But what riles readers is the nature of individual gifts. The database, however, described most gifts vaguely as food or beverage, travel, lodging or entertainment.

Armed with my results, I was able to approach lobbyists and lawmakers from a position of strength.

So the $900 flight on the corporate jet to the Super Bowl was simply listed as travel. I noticed, however, that the same date – the first Sunday in February – included $400 for entertainment and several gifts of food or beverage. Similar numbers showed up under the House majority leader's wife. The lawmakers involved acknowledged that it was a trip to the Super Bowl and the jet belonged to Monsanto, the St. Louis-based agribusiness giant.

Occasionally, however, I caught a break. In 2001, 10 lobbyists paid $5,400 to pay the tuition for a representative to attend a three-week seminar at Harvard. While some simply reported the payment as "other," half included a reference to tuition or the school.

Armed with my results, I was able to approach lobbyists and lawmakers from a position of strength. I usually knew more about the gifts than the lawmakers did. Because lawmakers often squirmed over the way their actions looked, connecting lobbyist gifts to actions is a difficult task.

The closest I came involved the Harvard seminar. Of the 10 lobbyists who paid for the legislator to attend, eight were insurance lobbyists. The lawmaker happened to be vice chairman of the House Insurance Committee. The year after his trip, he sponsored a House bill to ease regulatory oversight on property and casualty insurers and managed the debate on it.

A necessary, but time-consuming and often fruitless, task involved checking legislation, official journals and library clips to determine whether lawmakers helped lobbyists who gave them big gifts.

The most insightful result was from an interview with former Rep. Quincy Troupe, the recipient of $21,477 in gifts during his last two years in office.

Troupe, whose committee oversaw social services spending, described a symbiotic relationship between him and the Missouri Hospital Association. He said he protected the $1.3 billion in Medicaid funding for hospitals in the state budget as long as the hospital association's lobbyists would lobby others for his pet projects.

Most lawmakers deny any connection between gifts and legislation. The proof for any connection is usually circumstantial at best. That's because the connection is designed to be indirect.

The best description of a connection I found came from Anthony Nownes, a University of Tennessee political scientist who specializes in lobbyist behavior.

"Lobbyists' money buys something, and it appears to buy effort," Nownes said. "We tend to focus on votes. But most of the action is in writing the bills and during hearings. If I support a lobbyists' position, do I work harder to pass the bill? Or if I oppose the lobbyist, do I work a little less at trying to kill it?"

Contact Kit Wagar by e-mail at kwagar@kcstar.com.

Purged

continued from page 1

bloc that had been instrumental in getting Gov. Jeb Bush and the President elected.

Roughly 17 percent of Florida's population is Hispanic, the largest minority population in the state. It didn't seem logical that there would only be 61 on the list. After we double-checked our work to make sure we hadn't made a mistake, it left us with three possibilities to explore:

1. Are Hispanics less likely to commit felonies and less likely to register to vote? Neither turned out to be true. Hispanics make up about 11 percent of the state's registered voters, about 11 percent of Florida's prison population and 11 percent of the state's voting-age population. Given those numbers, it didn't make sense that less than 1 percent of the people on the state's purge list was Hispanic.

2. Were Hispanics hidden on the list under different race classifications? To answer this, we took a sample of 200 names from the list and examined them one by one, looking for Spanish surnames.

We settled on a sample of 200 after talking with Steve Doig, a journalism professor at Arizona State University and Matthew's father. The sample size gave us an acceptable error margin of just about 7 percent.

We obtained our sample by dividing 200 into 47,763, the total number of names on the list. We took the result, 239, and used Excel to number each name from 1 to 239 and repeated that process all the way through the list

Then we employed the highly scientific method of asking someone in the newsroom to shout out a number between 1 and 239. When we had our number, we used Excel's AutoFilter tool to display only those names with that number beside them.

We found a list of common Hispanic surnames on the Internet, and we enlisted a Hispanic colleague at the newspaper to help us identify people who could be Hispanic felons but listed under another race category.

We repeated this process three times to convince ourselves we hadn't made an error. Each time we determined that, at best, 3 percent of the people on the state's purge list were Hispanic. That fell far short of the 11 percent we expected to see.

3. Did Florida's election officials botch their list somehow? Given what happened during the 2000 presidential election, this was something we'd considered from the start. So when we ruled out the other potential explanations, we started investigating how the state had built their felon list so we could spot where a data flaw might have been introduced.

After seeing the story, *New York Times* reporter Ford Fessenden learned that there was more to the problem than ambiguity in race reporting.

Our Hispanic colleague helped again when he reminded us that Hispanics traditionally have trouble being placed into a particular racial or ethnic category. An official with the U.S. Census Bureau agreed with his observation.

The perfect example of this problem would be a man born in Cuba whose skin is black. That person might think of himself as Hispanic, and he could identify himself that way when registering to vote. But if someone else, such as a police officer, attempted to categorize him, he could easily be considered black.

We interviewed an official with Florida's Department of State who explained that they merged two data sets – one from the Florida Department of Law Enforcement and another from the Division of Elections – to create a list of people who were convicted felons registered to vote. Florida is one of a handful of states in which a felon's voting rights aren't automatically restored.

If you've used Microsoft Access to join two tables, you know an identifier common in both databases is needed to pull off the feat. And because the data might not be uniform in the tables, you also know what a headache it can be to make that match.

To ensure that John Smith the felon was also the same John Smith registered to vote, state officials used every identifier available to make their match. This included first and last names, date of birth, sex and race. If any of those fields didn't match across the tables, that person would not have been included in the purge list.

Because state officials used race for the matching process, we knew that's what had caused so many Hispanics to be left off the list.

We presented what we found to the state's elections chief, and she conceded that problems in classifying Hispanics into a single race category had probably resulted in their exclusion. Despite the admission, officials said they had no plans to fix the Hispanic felon flaw.

But apparently it's a lot harder to resist something published in *The New York Times*.

After seeing the *Herald-Tribune* story, *Times* reporter Ford Fessenden learned that there was more to the problem than ambiguity in race reporting. It turned out that the FDLE database didn't even offer Hispanic as a race category.

Later, the *Herald-Tribune* reported that prior to 1994, many supervisors of elections did not offer Hispanic as a race category. That virtually guaran-

teed that Hispanic felons wouldn't end up on the purge list. And it meant that Florida election officials had botched three felon purge lists since 1998.

Three days after the *Herald-Tribune* first reported the Hispanic voter flaw, Fessenden's story ran in *The New York Times*. Later that day, Florida Secretary of State Glenda Hood announced she was scrapping the list.

The flaw has cost the state millions of dollars and diminished its credibility, especially with Democrats and African-Americans who claim the list is designed to minimize their political power. The current purge list had only 9,500 Republicans, compared to about 28,000 Democrats.

The demise of the list could have a huge impact on the November presidential election, which polls show could be as close as the one in 2000. With another quick Excel analysis, we were able to illustrate the potential impact.

We began with the premise that the presidential election in Florida would be as close in 2004 as it was when it was decided by a 537-vote margin in 2000.

We wanted to see how many people on the purge list would have to vote along party lines to reverse the outcome of a hypothetical election that was as close as the Bush-Gore race.

What we found is that even if every Republican on the purge list voted for Bush, less than 36 percent of the Democrats would need to vote for Kerry to give Florida's electoral votes to Democrats. And if every Republican and Democrat on the list voted for their party's candidate, the outcome would be an 18,000-vote victory for the Democrats.

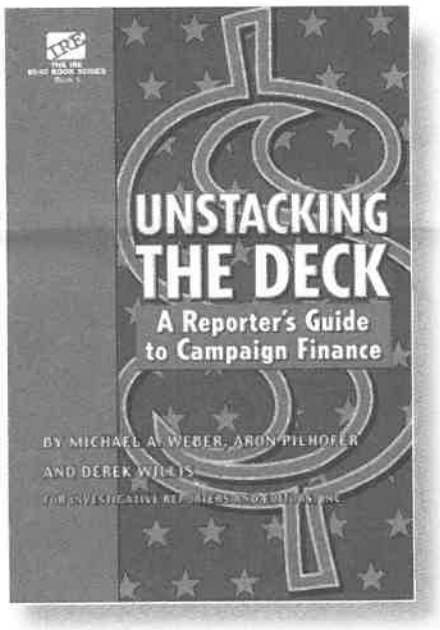
That little bit of math let us write this lead: "If the 2004 presidential election ends up being as close as the last one, Gov. Jeb Bush may have cost his brother his job."

Contact Chris Davis by e-mail at chris.davis@heraldtribune.com and Matthew Doig at matthew.doig@heraldtribune.com.

Unstacking the Deck:
A Reporter's Guide To Campaign Finance

Navigate the language and practices of campaign finance this election season with this guide from Investigative Reporters and Editors, Inc.

Journalists will find it invaluable for pursuing stories about the impact of money on elections, political parties and candidates at the federal, state and local levels.



- Understand the loopholes in soft money restrictions.
- Learn about the increased use of nonprofits to funnel money.
- Find out how to track where candidates spend their money.
- Learn how to find and read useful documents and data.
- Packed with resources, Web sites and story tips!

ORDER NOW

By phone: Call 573-882-3364 with your VISA, MasterCard or American Express.

By mail: Send your check to IRE, 138 Neff Annex, Missouri School of Journalism, Columbia, MO 65211

By Web: Visit www.ire.org/store to order online or download an order form.

IRE Members: $15 each • Nonmembers: $25 each

Plus postage: Media rate – $4 for the first book, $2 for each additional book

OPEN SOURCE

Tapping into MySQL's advanced features

By Aron Pilhofer, *Center for Public Integrity*

I have extolled the virtues of MySQL in *Uplink* because the program is insanely fast, rock-solid stable and free of cost. But over the past few months, I have discovered this open-source database server has a few more tricks up its sleeve.

In a couple of cases, these are features that are unavailable in Microsoft SQL Server 2000, which may be one reason to give MySQL another look even if you (like me) are working in an SQL Server newsroom.

Regular expressions

When you think of regular expressions you tend to think of a programming language like Perl. But MySQL supports an extremely advanced set of regular expression functions that make searching databases easier and more precise.

I recently investigated this feature when I was trying to find a way to efficiently join two tables of data by conditionally matching all or part of a company name in one field with campaign finance contributor information in another.

A simple wildcard search would have worked if not for a few confounding problems. First, I was looking for data on nearly 4,000 companies. Second, I was searching a campaign finance table of more than 4 million records. Last, I knew for a fact that often contributors will report only part of a company name. For example, "Kellogg, Brown and Root" will often appear as "Brown and Root" years after the Halliburton subsidiary officially changed its name.

I wanted a way to search for various combinations of "Kellogg, Brown and Root" in a single query, and that's exactly what regular expressions allowed me to do. Here is an example of what a search like this looks like in MySQL:

```
SELECT employer FROM test_table
where employer
REGEXP
'[[:<:]](^kellogg[[:space:]])?[[:<:]]Brown
(&|and) root'
```

The REGEXP operator can be used in the place of a LIKE clause in SQL, including in joins. While the regular expression itself might look intimidating to those who have not used one before, it's not that hard to pick up once you break it down into its component parts.

The above regular expression would read something like: Select any record from the employer field that contains the words "Brown" and "Root" separated by either the word "and" or an ampersand. Optionally, the word "Kellogg" may appear at the beginning of the record.

As you can imagine, regular expressions offer a great deal more precision over a wildcard search. But they also allow you to do certain things that a wildcard cannot do, for example match a ZIP code inside an address field. You could easily write a regular expression that would match a series of either five or nine numbers (optionally separated by a hyphen) that appear only at the end of an address string.

Those who are moderately familiar with Perl-style regular expressions won't have trouble figuring out the slight differences in MySQL syntax. A fuller explanation of regular expressions can be found in the online documentation at *http://dev.mysql.com/doc/mysql/en/Regexp.html*.

Group and group again

GROUP_CONCAT is a new function for MySQL as of version 4.1, and I have found it particularly useful for aggregating non-numeric data. Let me offer an example:

Say you have a number of ZIP codes and a number of towns within each ZIP, and you want to know in one query all the towns contained in each ZIP code. Something like this:

```
Zip     Town
55408   Culpepper Cove
55408   Smithtown
55408   Tarkentown
55409   Foremanville
55409   Yary Beach
```

A select query using group_concat would look like:

```
SELECT zip
GROUP_CONCAT(town)
FROM test_table
GROUP BY zip
```

And the result would look like this:

```
55408   Culpepper Cove, Smithtown,
Tarkentown
55409   Foremanville, Yary Beach
```

This is obviously a very simple example, but I have found GROUP_CONCAT to be particularly useful for analyzing campaign finance data and for cleaning federal contract data, where you often have odd companies showing up under various Dun & Bradstreet D-U-N-S numbers that uniquely identify the companies. It is by no means something I use every day, but for specific tasks it's another killer feature that keeps me coming back to MySQL.

Rollup your own

MySQL also includes a useful function that allows users to get not only aggregates but also subtotals. This function is a modifier to the GROUP BY clause, called ROLLUP. Expanding on the ZIP code example above, a ROLLUP of our sample data might look like this:

```
Town          Contribution        Party
Culpepper Cove $1,000
R
Culpepper Cove $900
D
```

```
Culpepper Cove $400
R
Smithtown              $500
        D
```

And our query might look like:

```
SELECT town, party,
sum(contribution) as Sum
FROM test_table
GROUP BY town, party WITH ROLLUP
```

This would result in the following:

```
Town            Party            Sum
Culpepper Cove D
$1,900
Culpepper Cove R                $400
Culpepper Cove NULL
$2,300
Smithtown               D
$1,100
Smithtown               NULL
$1,100
NULL                    NULL
$3,400
```

The records with the NULL value are subtotals, and the final record is a grand total. This is another feature I have found useful in analyzing campaign finance data, especially in calculating totals by ZIP code, town or contributor.

Conclusion

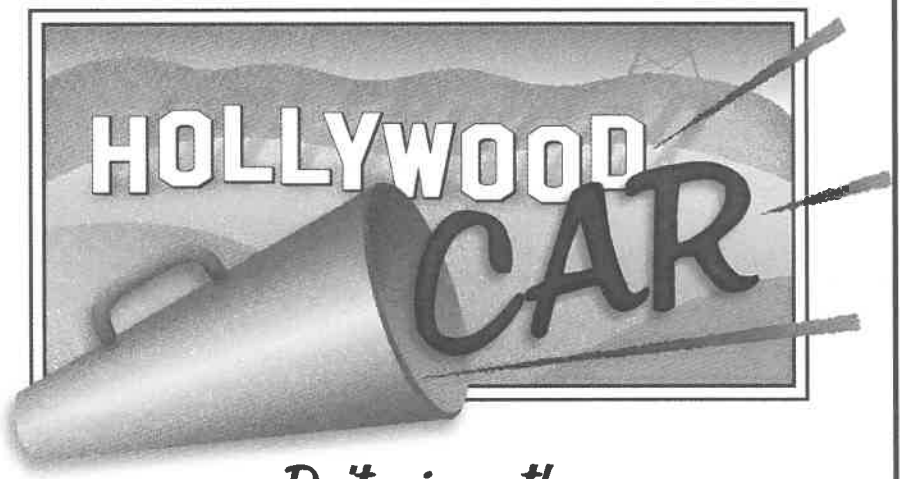
It is true that MySQL does not now, and probably never will, support the hundreds of features available in commercial database servers such as SQL Server. But right now, it includes 95 percent of the features CAR folks use 95 percent of the time, as well as a few features the other ones lack – and that makes MySQL a very intriguing alternative right now.

With version 5.0 (now in pre-beta testing), MySQL will add the last major feature, stored procedures. Even if it does not become your platform of choice for CAR work, it is worth investigating as an alternative to the high-priced database servers.

Contact Aron Pilhofer by e-mail at apilhofer@publicintegrity.org.

2005 Annual Computer-Assisted Reporting Conference

Hollywood, Calif. • March 17–20



Don't miss out!

Panels are planned on every newsroom beat and topic in the news. Key topic areas include local government, business, crime, education, health, environment and more. Issues will range from getting the most out of census numbers to following up campaign finance data to understanding federal contracts and international data. Panels will also focus on homeland security, military and infrastructure issues.

PLUS!

Exhibitors, networking, software demonstrations and special sessions for beginners, educators, broadcasters and advanced users.

For CAR beginners, select panels will be coordinated with specific hands-on classes. The combination will give particiapnts a mini-boot camp experience with plenty of opportunity to apply what they learn.

Hands on classes will cover basic to advanced skills in spreadsheets, database managers, using the Internet, building your own database, mapping, statistics, social network analysis, SQL and more.

Host:

KNBC-Los Angeles

Conference Hotel:

Renaissance Hollywood Hotel • 1755 N. Highland Ave. • Hollywood, CA 90028

Hotel reservations: Call 800-468-3571 by Friday, Feb. 18, to get the discounted room rate of $164 plus tax. When making your reservation please ask for the IRE/2005 Annual CAR Conference room block.

Presented by Investigative Reporters and Editors, Inc. and the National Institute for Computer-Assisted Reporting

For more information, registration and the latest schedule, visit

www.ire.org/training/hollywood05/

Tech tip...

Scaping for data the Perl way

By John Perry, *The Oklahoman*

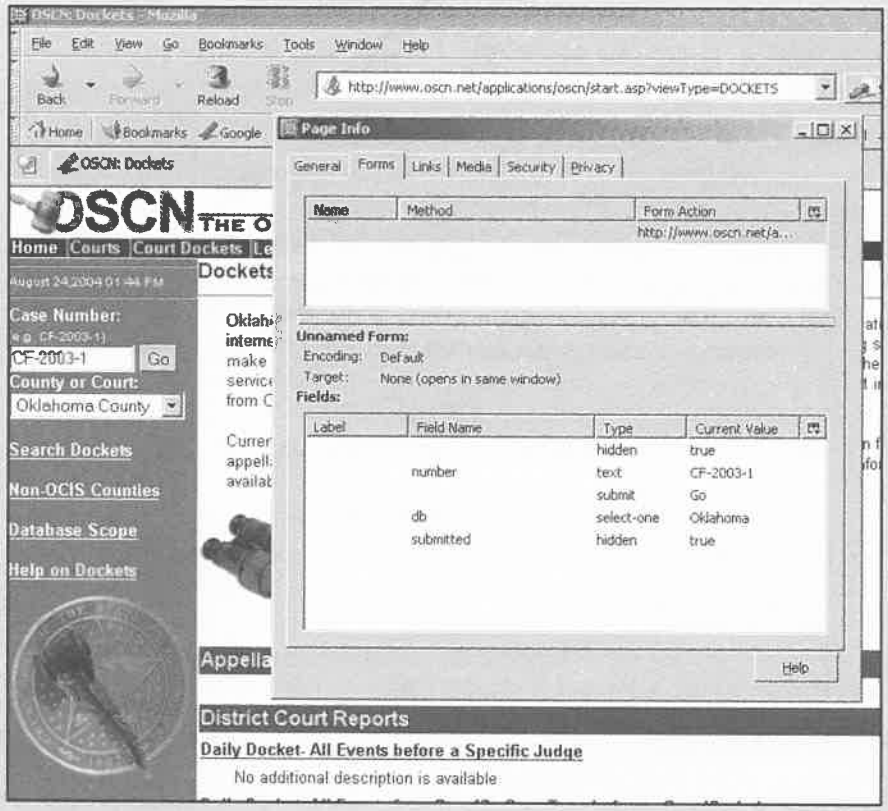
In the March/April edition of *Uplink*, Derek Willis wrote about using the free utility Curl to automate the task of downloading pages from Web sites and online database search applications. Sometimes you need a smarter tool that cannot only find Web pages, but also looks at the content and decides what to do with it.

That's where Perl – and other scripting languages such as PHP and Python – come to the rescue. I prefer Perl because, first of all, it's what I know. My philosophy is the solution that doesn't require learning a new technology is always the best.

I also believe that Perl offers some unique tools that make the job of Web spidering – that is, grabbing entire files or Web pages – and scraping so much easier. Thousands of modules, pre-written scripts that provide functions to simplify programming tasks, are available to Perl users. Two of them, WWW-Mechanize and HTML-TokeParser, are godsends for Web scrapers. (For more about Web scraping and the IRE and NICAR Scraping Project, see page 3).

WWW-Mechanize acts like a virtual Web browser and hides the technical intricacies of HTTP requests and responses behind simple-to-use functions such as "follow_link," "back" and "click."

HTML-TokeParser is one of several tools Perl provides for pulling useful information from Web pages. The vagaries of poorly written HTML code and the random intrusion of new lines make searching for information with regular expressions a nightmare. TokeParser turns a Web page into a stream of tokens, those bits of information that can be start tags, end tags, comments or text. You pull tokens out of the stream one-by-one, looking for patterns that mark the location of desired information, or you can jump ahead to find specific tags with particular attributes – no regular expressions required.



Web scraping came in handy during a recent investigation of the Oklahoma County District Court and the court clerk's office. We received a tip from the local county courthouse that an unusual number of bonds that should have been forfeited when a criminal defendant skipped a court date were being exonerated. To get an idea of what was going on, we wanted to look at all Oklahoma County cases in which the bond had been forfeited.

The Oklahoma Supreme Court, which maintains the docket data for Oklahoma's eight biggest counties, is generally responsive to data requests. But they are also overworked and understaffed and in the middle of updating their software. The data you get is complex – spanning 40 tables – and the docket text file is a nightmare to parse. Their Web search result pages, by contrast, are simple and seem designed for easy parsing.

So to speed things up, we requested only a list of bond forfeiture case numbers. With a Perl script using the

Mechanize module, I grabbed the docket info page for each of the cases and saved them in a directory. I then wrote a script that used the TokeParser module to search for key words in the docket text.

We later found that the local court clerk's office was less than diligent in properly coding all the bond forfeiture cases, so I combined the two scripts to search through all of the past year's Oklahoma County criminal cases online to find those that mentioned bond forfeitures in their docket text.

The first step in Web scraping is to find the information about the search page that you'll need for the purpose of submitting searches. And for this task, the Mozilla Web browser is a great help. The page information function (View->Page Info) gives you a list of forms on a page, the fields in each form and the information they contain. You can fill out the search form, look at the page information window and see which fields you have to set in your script.

So to find and save all the cases in our list of bond forfeitures, we had to accomplish these tasks: get the search form page; set the "db" field to "Oklahoma" for Oklahoma County and the "number" field to the case number; click the search button; save the page contents to a file; then go back to the search form for the next case. The Perl code looks like this:

```
use strict;use warnings;
use WWW::Mechanize;

# Get case numbers (starting
with CF, CM or TR) from file
open CASES, 'c:/bonds/
qbondforft.txt';

# Pull the case numbers into an
array
chomp( my @cases = grep /
^(CF|CM)/, <CASES> );
close CASES;

# create browser object and get
the page
my $ua = WWW::Mechanize->new();
$ua->get('http://www.oscn.net/
applications/oscn/
start.asp?viewType=DOCKETS');

# Step through the case num-
bers, saving content to file
named with case number
for my $case ( @cases ) {
    open OUT, ">c:/bonds/pages/
$case.html";

    #set values in the form and
submit
    $ua->set_fields( db =>
'Oklahoma', number => "$case"
);
    $ua->click;


    #save content to file, wait
a second, then return to search
form
    print OUT $ua->content;
    close OUT;
    sleep(1);
    $ua->back;
}
```

The first line adds the "strict" and "warnings" directives called pragmas to your code, which tells Perl to give you more information about anything that might go wrong. The next "use" statement adds the Mechanize module to your script. The open statement creates a file handle, CASES, that gives access to the file listed in the statement.

The statement on line eight uses the grep function to pull all of the text lines from the file handle CASES that match the regular expression /^(CF|CM)/, which will include lines starting with CF (criminal felonies) or CM (criminal misdemeanors), and save them in the array, @cases. It then chomps the array, removing all the return and line feed characters off the end of the array elements.

Line 12 is an example of Perl's object syntax. Without getting mired in object-oriented programming jargon, the statement uses the "new" method to create a Web browser object and save it in a variable called $ua. We can do things with our browser by applying methods to it using the arrow operator. The statement on the next line, for example, uses the "get" method to grab the search form page and save all the information about it in $ua.

In lines 16 through 28 we use a "for" loop to step through the case numbers saved in the @cases array, open a new file using the case number as the file name, set the fields on the search page with the "set_fields" method, click the search button, and print the content of the resulting page to the open file handle.

We then close the file, pause for a second with the "sleep" statement, and then go back to the search page by means of the "back" method. The one-second sleep is necessary because, like most Web servers, the Supreme Court's server protects itself from denial of service attacks by temporarily blocking requests from sources that make too many requests over a period of time.

continued on page 18

Tech tip...

continued from page 17

Now that we have the pages and we can search them:

```
use HTML::TokeParser;

opendir DIR, 'c:/bonds/pages';
my @files = grep /C(M|F).+html/
readdir( DIR );
close DIR;

for my $file ( @files ) {
    my $p = HTML::TokeParser-
>new( "c:/bonds/pages/$file" );

    while( my $toke = $p-
>get_token ) {
        if( $toke->[0] eq 'C'
&& $toke->[1] =~ /
DocketID:(\d+)/ ) {
            $p->get_tag( 'tr'
);
            $p->get_tag( 'tr'
);
            $p->get_tag( 'td'
);
            my $text = $p-
>get_trimmed_text( '/tr' );
            if( $text =~ /
(VACATE|ASIDE|EXONERATE)/ ) {
                print "$file:
$text\n\n" if;
            }
        }
    }
}
```

Again, the "use" statement on the first line imports the helpful module, in this case HTML-TokeParser.

The "opendir" statement opens a directory handle, "DIR," that gives access to the directory where the pages were saved. On line three, the "readdir" function gives us a list of files, and the "grep" function pulls out the file names that start with "CM" or "CF" and end with "html" and saves them in the array @files.

Next we step through the file list with a "for" loop. On line eight, we create a new token stream by applying the "HTML::TokeParser->new" method to the file. Then we start searching for the docket text by stepping through the tokens one by one with the while loop that begins on line 10.

The docket entry source code look like this:

```
<!–DocketID:41280303–>
<TR><TD><FONT FACE=ARIAL SIZE=2
COLOR=black>12/16/2003</FONT></
TD>
<TD><FONT FACE=ARIAL SIZE=2
COLOR=black>TEXT</FONT></TD>
<TD><FONT FACE=ARIAL SIZE=2
COLOR=black>-</FONT></TD>
<TD><FONT FACE=ARIAL SIZE=1
COLOR=black>Johnson, Kizzy</
FONT></TD>
<TD><FONT FACE=ARIAL SIZE=1
COLOR=black>41280303</FONT></
TD>
<TD><FONT FACE=ARIAL SIZE=1
COLOR=black>Dec 16 2003
3:17:51:813PM</FONT></TD>
<TD><FONT FACE=ARIAL SIZE=2
COLOR=BLACK>-</FONT></TD><TD
ALIGN=RIGHT><FONT FACE=ARIAL
SIZE=2 COLOR=GREEN>$ 0.00</
FONT></TD></TR><TR><TD><FONT
FACE=ARIAL SIZE=2
COLOR=> </TD>
<TD COLSPAN=8><FONT FACE=ARIAL
SIZE=2 COLOR=black>THE STATUS
OF THE BOND ENTRY DETAILED IN
DOCKET SERIAL #39254233 ABOVE
HAS CHANGED TO READ AS FOLLOWS:
<BR>SURETY BOND FOR JOHNSON,
KIZZY POSTED BY BURCH, MICHELLE
(SURETY:NATIONAL AMERICAN )
(POWER NUMBER:NA5-793071  2625-
0437 ), IN THE AMOUNT OF
$5,000.00, POSTED 02/21/2003,
EXONORATED 12/05/2003
```

They are conveniently marked with a comment, "<!--DocketID:41280303-->" that we can't see when viewing the page in a browser. The text we're looking for is stored in cells of an HTML table. The rows are set off by the <tr> and </tr> tags. The cells in each row are set off by the <td> and </td> tags.

To get to the docket text, we start at the comment, then step down two rows and one cell with the three "get_tag" statements on lines 12-14. On line 15, the "get_trimmed_text" method grabs all the text up to the next row end tag, </tr>, eliminating any extra white space. If the docket text matches any of the words in the regular expression on line 16, it prints the text with the file name.

This is a relatively simple example of the capabilities of Perl and its collection of Web scraping tools. To delve more deeply into scraping, "Spidering Hacks" by Kevin Hemenway and Tara Calishain and "Data Munging with Perl" by David Cross are very helpful. You can also find tutorials on using WWW-Mechanize and HTML-TokeParser at *www.perl.com* and *www.perlmonks.org*. You can find the documentation for these and other Perl modules at *http://search.cpan.com*.

Contact John Perry by e-mail at JPerry@Oklahoman.com.

ELECTIONS

Mapping candidate support

While reporting on presidential campaign fundraising, Ron Campbell of *The Orange County Register* used Federal Election Commission campaign contribution data and ArcView 8.3 to show where President Bush and Democratic Party challenger John Kerry were drawing financial support.

Campbell prepared these maps (Bush on top left, Kerry on top right) showing the locations of individual donors who gave the maximum $2,000 contribution to each candidate. The maps show that Bush was successful in attracting contributions from a wide area. Kerry's big financial support, in contrast, has come mostly from metropolitan areas.

For both candidates, Campbell queried FEC data of individual contributions to count all $2,000 donations by five-digit ZIP code. He then joined those tables to a ZIP code map in ArcView and created a national map (bottom left) and one for southern California (bottom right).

Starting

continued from page 1

nizations, the thought is intimidating. They know that with CAR, as with most skills, it's use it or lose it.

The real trick is changing that mind set. When working in a small- to medium-sized newsroom, you must not use size as an excuse, even if you have fewer CAR converts and resources than your counterparts in larger newsrooms.

Easier said than done.

But journalists in some smaller newsrooms are doing outstanding CAR work. A good place to see this in action is in *Uplink*'s First Venture feature, which provides useful lessons for journalists in smaller newsrooms. A recent example: Dave Moore of the *Columbia* (Mo.) *Daily Tribune* wrote about his analysis of racial disparity in juvenile arrests in the May/June 2004 issue. When Dave got the records on a CD, they were a mess. He could have given up. But with some help from NICAR, he got the data straightened out and a good story into the paper.

When I worked in a smaller newsroom, I had some rules to keep on track. What follows are 10 of those dos and 10 don'ts. Cut out the ones that make sense to you and paste them on your computer.

Do:

1. Have reporters sit next to you while you do your analysis. They might pick up some skills, and you'll get an ally who will see just how much you add to your newsroom. No Lone Rangers - share knowledge. Keep a copy of Brant Houston's "Computer-Assisted Reporting: A Practical Guide" on hand so you can run through a few exercises with other reporters and help them reinforce their skills.

2. Insist your reporters gather all documents electronically if possible. City and school board budgets are most likely prepared in a spreadsheet, such as Microsoft Excel. Having Excel on your computer won't help your cause if you don't have data to analyze.

3. Attend news meetings at which brainstorming for story ideas happens. Wait for your opening, and when you see the possibility of how CAR will improve a story, pounce. Follow through.

4. Get to be the "numbers person" in your newsroom. Sometimes a single number or the comparison of just two numbers makes a story work. Such a story will garner unbelievable praise from editors, who won't know how easily you dashed it off. Trust me. All experienced CAR-hands have an example of how editors gushed over their contribution to a story by just adding a number for context. Cost to you: maybe 10 minutes. Not sure where to start with the math? Get a copy of Sarah Cohen's "Numbers in the Newsroom" in the IRE Beat Book Series.

5. Start small. Editors might like big packages, but they really adore daily copy that makes their lives easier. Show them how CAR can add context and reader interest. It works even better when you can do it on a deadline story. They'll look good and may be willing to support your bigger efforts.

6. Broaden your definition of CAR. It's more than sorting in Excel and grouping in Access. Keep learning new techniques. If your local university lacks a good journalism program, check out the sociology and political science departments. You might find a quantitative data analysis class (academic speak) or a social statistics class you can audit. Keep pressing yourself to learn how to apply your learning.

7. Get to know your graphic artists. Take them to lunch. Let them know that how they present your analysis means a lot to you and that you appreciate what they do. Once you build rapport with them, persuade them they need to take a CAR workshop to improve their information graphics.

8. Bookmark 10 Web sites you can use to find data or techniques or answers to questions on deadline. Some of those sites should be lists of tried-and-true links gathered by journalists for journalists. One of your bookmarks should be the NICAR 'Net Tour at *www.ire.org/training/nettour*. Spend time getting familiar with the 'Net Tour so you know your way around it on deadline.

9. Carve out a neglected niche in your newsroom that is CAR-intensive. For example, if you have no one dedicated to analyzing Census releases, master them and pitch them at reporters whose beats overlap the data. Another example: Learn how to analyze crime stats, especially if your cops and court reporters don't know CAR. That will help add depth and context to the cops and court stories while building your skill. A good place to start is "Understanding Crime Statistics" by Kurt Silver in the IRE Beat Book Series.

10. Stay passionate and love what you do. Remember that you're the one on the cutting edge. You're the one building what will not only help safeguard your career, but will also serve your community by adding more depth and context to your stories.

Don't:

1. Take no for an answer when asking for hardware or software. We don't take no from difficult sources. We shouldn't from our editors either. If there is software you want, be persistent. Find out when your editors figure out next year's budget. Get added to it. If you must, embarrass your editors by showing them what CAR resources your competition has or what other news organizations your size are doing with CAR.

2. Be a recluse. Keep networking with CAR folks around the country. Go to the annual Investigative Reporters and Editors CAR Conference March 17-20 in Hollywood, Calif. (See *www.ire.org/training/hollywood05* for details). Start planting the seed now with your editors and tell them about the payoff if you attend. Find people who are doing CAR-based stories that you admire. Bug them. Ask if you can contact them for advice. Ask them how they learned. Get on the

2005 CAR Training

These unique seminars, taught by IRE and NICAR's experts, train journalists to acquire electronic information, use spreadsheets and databases to analyze the information and to translate that information into high-impact stories. In addition, IRE and NICAR provide follow-up help when participants return to their news organizations.

Mapping Data for News Stories:
Columbia, Mo.
Learn how to uncover news stories by mapping data with geographic information system software.
- Jan. 14-16
- Aug. 19-21

Boot Camps:
Columbia, Mo.
- Jan. 9-14
- March 20-25
- May 15-20
- Aug. 7-12

Boot Camp for Editors:
Columbia, Mo.
Tailored to the needs of newsroom managers.
- April 8-10

Advanced CAR Statistics:
Tempe, Ariz. (at ASU)
Aimed at strengthening the skills of reporters who want to move beyond basic CAR.
- Feb 18-20

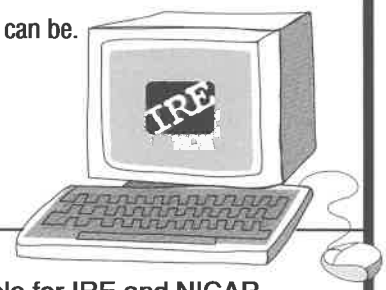
For more information: **www.ire.org/training/bootcamps.html**

What previous Boot Campers have said about the experience:

"They were able to talk in a real way about issues that reporters face when practicing computer-assisted reporting skills ..."
– Analisa Nazareno, *San Antonio Express-News*

"The intense, small group workshop helped me understand how powerful database research can be. ... I now consider CAR one of the most powerful tools to expose wrongdoing."
– Holly Whisenhunt, WOAI-San Antonio

"I realized I hadn't been fully tapping my beats, and you don't have to be a computer nerd to do computer-assisted reporting."
– Teresa Taylor Williams, *Muskegon Chronicle*



Fellowships

A limited number of minority and small-news organization fellowships are available for IRE and NICAR workshops. Visit **www.ire.org/training/fellowships.html** for more information and an application.

NICAR-L listserv. Go to *www.ire.org/membership/subscribe/nicar-l.html* to learn how. You'll have a responsive group of CAR experts who most likely have faced the same problems you're facing.

3. Let public officials win when they refuse to give you data digitally. Learn some good responses to the hurdles they throw in your face. Know the open records law in your state. Better yet, in this day of shrinking civil liberties, become a champion for open records in your community.

4. Stay in a newsroom that has absolutely NO appreciation for CAR. They're getting rarer, but they're still not extinct. If you are taking charge of your career and building your CAR skills, you'll be marketable. You are investing in a passion. You can't build where there's no foundation.

5. Forget the IRE Resource Center when trolling for CAR ideas. Search for "computer-assisted reporting" in the 20,000-story database, and as of September 2004, you'd get 487 hits. Use them to stimulate your imagination. Go to the IRE home page each day and click on Extra! Extra! to keep up with the latest CAR work available on the Web.

6. Get totally caught up with the computer and forget about stories unless you really have plans to be an IT professional. The computer is a tool. You're working on a way to approach journalism. It's been more than 30 years since Phil Meyer published the first edition of "Precision Journalism." CAR is as much about scientific method as it is the computer. If this doesn't yet click, order the fourth edition of "Precision Journalism" from the IRE bookstore and see what it's about.

7. Be afraid to think big. When I worked at the *Savannah Morning News* (with about 75,000 circulation on Sundays), we accomplished more by not reminding ourselves of our size but thinking about what we could do. We never used our size as an excuse. We were one of the smallest news organizations in the country that regularly conducted its own public opinion surveys. We learned how to do it regardless of our size.

8. Fall for Stockholm Syndrome in your newsroom. Editors who don't support or even appreciate what you're trying to accomplish do not deserve your empathy. They are not your friends.

9. Try to make everyone in your newsroom a CAR believer. You'll spread yourself too thin. Look for converts and work with those who "get it" or show an interest in learning. Build your program one reporter at a time.

10. Get discouraged. You're doing work vital to journalism and those you serve. To paraphrase newspaper editor and sociologist Robert Park, computer-assisted reporting is like art to an artist, less a career than a form of excitement and a way of life.

Contact David Donald, the national training director for Investigative Reporters and Editors at ddonald@ire.org.

readme.txt

Some of the best journalists in the business recently spoke at the IRE Annual Conference in Atlanta about putting together a computer-assisted reporting project from start to finish. You can purchase tapes of these panels and related tipsheets.

Order tapes from Sound Images Inc. by providing the Session ID number. Contact Sound Images at 888-649-1118 or service@soundimages.net. A complete list of the available sessions can be found at *www.ire.org/store/tapes/atlanta04.pdf.*

IRE members can search for the tipsheets online and view them online at *www.ire.org/resourcecenter/initial-search-tipsheets.html.* You may also order the tipsheets by contacting the IRE Resource Center at 573-882-3364 or rescntr@ire.org and providing the tipsheet number.

"CAR from start to finish: Choosing stories." John Erickson of the *Dayton Daily News* and Douglas Pardue of *The* (Charleston, S.C.) *Post and Courier.* (Session ID No. IRE04-009). Pardue provides advice on looking for cues to good stories in Tipsheet No. 2065. Erickson's Tipsheet No. 2001 also provides ideas for identifying stories while on the beat.

"CAR from start to finish: Fighting for data." Russell Carollo of the *Dayton Daily News* and Jennifer LaFleur of *The Dallas Morning News* (Session ID No. IRE04-014). LaFleur's Tipsheet No. 2095 shows how data restrictions have hampered reporting.

"CAR from start to finish: The data is in – now what?" Ted Mellnik of *The Charlotte Observer* and David Milliron of *The Atlanta Journal-Constitution.* (Session ID No. IRE04-019). Tipsheet No. 2132 offers advice about looking at data before analyzing it.

"CAR from start to finish: Putting it all together." Rose Ciotta of *The Philadelphia Inquirer*, Shawn McIntosh of *The Atlanta Journal-Constitution* and George Papajohn of the *Chicago Tribune.* (Session ID No. IRE04-024). The panelists share their advice in Tipsheet No. 2099.

For an overview of the lessons learned in two decades of CAR, order "Blending traditional reporting with data analysis: What we've learned in 20 years" (Session ID No. IRE04-099) with Sarah Cohen of *The Washington Post*, John Erickson of the *Dayton Daily News*, and Deborah Nelson of the *Los Angeles Times.* Nelson has a checklist to use at the start of a CAR project (Tipsheet No. 2092). Erickson provides guidelines for editing the CAR project (Tipsheet No. 2087). Cohen shares advice for working in teams (Tipsheet No. 2128).

The IRE Resource Center also has tipsheets from earlier conferences:

Mark Nichols of *The Indianapolis Star* provides ideas on how editors can help guide the CAR story. (Tipsheet No. 2049)

Barbara Vobejda of *The Washington Post* provides advice for putting together the CAR project. (Tipsheet No. 1945)

Debbie Wolfe of the *St. Petersburg Times* provides a quick checklist of questions for editors so they can be CAR-savvy through every step of the reporting process. (Tipsheet No. 1828)

Griff Palmer of the *San Jose Mercury News* offers tips about editing and organizing a large data-heavy project, from getting a strong start and helping it grow to making it relevant for readers. (Tipsheet No. 1557)

Guiding

continued from page 1

Ullmann liked two-person investigative teams because he could partner people with different skills, and they would finish more quickly.

Ullmann said good stories come from good editors who are well-read and good reporters working out in the field. While data is a valuable resource for finding story ideas, it's just a starting point, he said. Reporters should not be too proud of themselves when they get the data because the data is rarely the story. Journalists are in the business of storytelling, not describing numbers. Stories are always about victims and protectors, and it's important to look beyond the data to find them.

When a reporter pitches a prospective story, Ullmann said an editor should ask two questions. First, what is the best-case scenario of the story? Second, would the reporter read the story if his or her name wasn't on it?

Reporting goals

Once a news organization approves a computer-assisted project and the reporting begins, Ullmann said it's important to give meaning to the data. That means the reporters need to find people who can reflect what the data shows and identify experts who can help interpret conclusions.

Ullmann recommended transcribing soon after the interview. He said reporters should throw out any notes they took during the interview because these are not as accurate as the transcribed tape.

Numbers can mean many things; Ullmann said it's important to include numbers that can support other interpretations in the story. He suggested working just as hard to disprove a story's thesis. Show the findings to the "enemy" early in the reporting to avoid surprises after publication or broadcast.

Reporters and editors should meet fre-

quently for story conferences. When he was an editor, Ullmann said he barred computer printouts and notebooks from these meetings so numbers would not rule reporters.

Reporters should never wait until they finish reporting before starting to write, Ullmann said. He liked to see drafts from his reporters every few weeks. Reporters should write what they know as they report. This strategy allows the news organization to publish or broadcast earlier than planned if necessary. In addition, it gets the reporters to focus on the story, not on each individual document or source. Ullmann suggested that reporters distribute their drafts to editors around the newsroom to find out if they get excited about the story.

The editing

Ullmann said it's important to allow substantial time for editing and rewriting. That may involve shortening the story, but not always. Ullmann said long stories work if readers and viewers care about them.

Charts, graphics and pictures are one way to increase a story's pace by cutting number-heavy material out. Ullmann said editors should include the photo and graphic staff from the beginning. They are partners in the presentation of the project, and they are critical to making it work.

When the project is ready for publication or broadcast, Ullmann stressed one final step: line-by-line editing. Reporters should do this with their data. When Ullmann supervised projects, he wanted to see the source of every number in the story, whether it was in a document or came from a formula. He asked his reporters to redo calculations in front of him to make sure they were correct.

If editors follow these steps, Ullmann said they would not be stuck on the bench; instead they will become valuable members of the team.

Contact Tim Ragones, a former analyst in the IRE and NICAR Data Library, by e-mail at Tim.Ragones@wsaz.com.

IRE and NICAR Services

Investigative Reporters and Editors, Inc. is a grassroots nonprofit organization dedicated to improving the quality of investigative reporting within the field of journalism. IRE was formed in 1975 with the intent of creating a networking tool and a forum in which journalists from across the country could raise questions and exchange ideas. IRE provides educational services to reporters, editors and others interested in investigative reporting and works to maintain high professional standards.

Programs and Services

IRE Resource Center: A rich reserve of print and broadcast stories, tipsheets and guides to help you start and complete the best work of your career. This unique library is the starting point of any piece you're working on. You can search through abstracts of more than 19,000 investigative reporting stories through our Web site.
Contact: Jaimi Dowdell, jaimi@ire.org, 573-882-3364

Database Library: Administered by IRE and the National Institute for Computer-Assisted Reporting. The library has copies of many government databases, and makes them available to news organizations at or below actual cost. Analysis services are available on these databases, as is help in deciphering records you obtain yourself.
Contact: Jeff Porter, jeff@ire.org, 573-882-1982

Campaign Finance Information Center: Administered by IRE and the National Institute for Computer-Assisted Reporting. It's dedicated to helping journalists uncover the campaign money trail. State campaign finance data is collected from across the nation, cleaned and made available to journalists. A search engine allows reporters to track political cash flow across several states in federal and state races.
Contact: Brant Houston, brant@ire.org, 573-882-2042

On-the-Road Training: As a top promoter of journalism education, IRE offers loads of training opportunities throughout the year. Possibilities range from national conferences and regional workshops to weeklong boot camps and on-site newsroom training. Costs are on a sliding scale and fellowships are available to many of the events.
Contact: David Donald, ddonald@ire.org, 573-882-2042

Publications

The IRE Journal: Published six times a year. Contains journalist profiles, how-to stories, reviews, investigative ideas and backgrounding tips. *The Journal* also provides members with the latest news on upcoming events and training opportunities from IRE and NICAR.
Contact: Len Bruzzese, len@ire.org, 573-882-2042

Uplink: Bimonthly newsletter by IRE and NICAR on computer-assisted reporting. Often, *Uplink* stories are written after reporters have had particular success using data to investigate stories. The columns include valuable information on advanced database techniques as well as success stories written by newly trained CAR reporters.
Contact: David Herzog, dherzog@ire.org, 573-882-2127

Reporter.org: A collection of Web-based resources for journalists, journalism educators and others. Discounted Web hosting and services such as mailing list management and site development are provided to other nonprofit journalism organizations.
Contact: Ted Peterson, ted@ire.org, 573-884-7321

For information on:

Advertising: Pia Christensen, pia@ire.org, 573-884-2175

Membership and subscriptions: John Green, jgreen@ire.org, 573-882-2772

Conferences and Boot Camps: Ev Ruch-Graham, ev@ire.org, 573-882-8969

Listservs: Ted Peterson, ted@ire.org, 573-884-7321

Mailing Address:
IRE, 138 Neff Annex, Missouri School of Journalism, Columbia, MO 65211

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

NON-PROFIT ORG.
U.S. POSTAGE
PAID
Jefferson City, MO.
Permit NO. 89

Ira Chinoy
Philip Merrill College of Journalism
1117 Journalism BLDG
University of Maryland
College Park MD 20742

Uplink

Uplink Info

A newsletter of the National Institute for Computer-Assisted Reporting

Editor
Brant Houston
brant@ire.org

Managing Editor
David Herzog
dherzog@nicar.org

Asst. Managing Editor
Jeff Porter
jeff@nicar.org

Art Director
Lisa Triefenbach

Copy Editor
Pia Christensen

Contributing Editors
Megan Clarke
Andrea Lorenz
Brian M. Hamman
Catherine Rentz Pernot

NICAR is a joint program of Investigative Reporters and Editors Inc. and the Missouri School of Journalism.

NICAR services include supplying journalists with government databases, training programs, tipsheets and data analysis.

Editorial
573-884-7711

Subscriptions
573-882-2772

Director of Publications
Len Bruzzese
len@ire.org

Advertising Coordinator
Pia Christensen
pia@ire.org

Subscription Administrator
John Green
jgreen@ire.org

Subscriptions
IRE members $40, nonmembers $60

Uplink Address:
IRE-NICAR, 138 Neff Annex
Missouri School of Journalism
Columbia, MO 65211

Postmaster: Please send address changes to IRE-NICAR.