Uplink

March/April
2001

ELECTIONS

CAR in the aftermath

By Geoff Dougherty
The Miami Herald

On the morning of Nov. 8, I sat in *The Miami Herald*'s cafeteria with three colleagues. Bleary eyed from watching Florida's presidential election returns roll in until 3 a.m., we talked over the results. Five hundred votes, in a state known for ballot fraud, would determine the presidency.

It seemed clear: We needed to take a comprehensive look at the Florida election. Who had voted? Did they do it legally? How had their ballots been counted (or discarded)? Were there signs of organized fraud?

Answering those questions would take us thousands of hours over the next four months. Over that time, we had cranked up a data collection and analysis effort of unprecedented scope.

Databases galore

We've gathered and analyzed more than 50 databases, including:

• Information from the Florida Secretary of State's office on each registered voter in Florida, and whether they cast a ballot on Nov. 7.

• Precinct-level election result files for all of the more than 5,000 precincts in Florida, including information on the number of undervotes and overvotes. These came in both electronic and paper format (that was then hand-entered) from county election supervisors.

•Databases from the court and police agencies in Dade, Broward and Palm Beach counties, identifying potential voters who were logged into jail as illegal immigrants, people who had vouched for their non-citizen status on jury selection forms and people declared incompetent to vote by the courts.

• A database from the Florida Department of Corrections identifying millions of people who had served state prison time after being convicted of felonies.

• A database of all convicted felons who had been granted clemency and had their right to vote restored by the Florida Office of Executive Clemency.

• A database from the Florida Department of Health of each person who had died in the past year in Miami-Dade County, along with databases from the Florida Secretary of State identifying voters who were targeted for removal from the rolls. They were being re-

Continued on page fourteen

FIRST VENTURE

Everday methods

By Joe Ellis
KMOL-TV, San Antonio

I thought I'd be eager to get back to my computer immediately after bootcamp, but that was before bootcamp. I remember my brain feeling like it had run a week's worth of PT. And it had...following the cadence of Sgt. Tom McGinty and company for six days, 10 hours a day. It showed me just how out of shape my mind really was.

I took what I felt was a well-deserved weekend away from my thoughts of work, news and computer-assisted reporting. But it wasn't long before I was finding ways to use my newfound skill.

Since then, CAR has become an everyday method for me.

When I'm looking for information, among the first things I ask the "Public Interference Official" is "how are these records kept?" "In what form are they

Continued on page twenty

Inside Uplink

Reporters used many computer-assisted reporting techniques to delve into the post-election quagmire. Former *Miami Herald* reporter Steve Doig, now an Arizona State University professor, devised a method to apportion uncounted Florida votes to the candidates. Nancy Cook Lauer of *The Tallahassee Democrat* found that lower income Florida counties had more uncounted votes than their affluent counterparts. A roundup of several election stories is also included in this issue.

First Ventures
Michele Rubenstein of WMAQ-TV, Chicago discusses her project on infections in hospitals.
See page ten

Ammonia leaks
Dawn Fallik of the *St. Louis Post-Dispatch* talks about databases that can provide a twist to covering the methamphetamine story.
See page twelve

Census
Paul Overberg of *USA Today* shares tips to cover the census and gives readers 10 story ideas.
See page sixteen

Uplink

March/April 2001

Volume 13, Number 2
A newsletter of the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston

MANAGING EDITOR
Mary Jo Sylwester

ASSOCIATE EDITORS
Amy Sherrill
Mike Sherry

ART DIRECTORS
Kerrie Kirtland
Yan Jin

NICAR is a joint effort of Investigative Reporters and Editors and the University of Missouri School of Journalism.

NICAR services include hands-on newsroom training in computer-assisted reporting, special academic and advanced training in data analysis.

DIRECTOR OF PUBLICATIONS
Len Bruzzese

SUBSCRIPTION ADMINISTRATOR
John Green

Uplink is published bimonthly by the National Institute for Computer-Assisted Reporting, 138 Neff Hall Annex Columbia, MO 65211. (573) 882-0684. Subscriptions are $40 for IRE members, $60 for nonmembers.

Postmaster: Please send address changes to NICAR. Send e-mail to jgreen@nicar.org

ANNOUNCEMENT

Uplink improvements

By MaryJo Sylwester
IRE and NICAR

For more than a decade, *Uplink* has brought you inspiring tales, technical how-to's, story ideas, tips for finding information and more. The newsletter has chronicled the evolving realm of computer-assisted reporting and offered support and guidance to hundreds of journalists worldwide.

Now we'd like to do more.

Starting with this issue, we will publish *Uplink* on a bi-monthly schedule (every other month). In the next issue, you'll see even more content, with at least 28 pages and supplementary materials on the Web.

Each issue will have a central theme – i.e. education in one issue, transportation in another – giving journalists not only the inspiring tales, but also the means to pull them off in their own community. There will be sidebars listing databases and where to find them; Web sites and phone numbers to sources; and listings of other stories on the same or similar topics.

In addition to the theme stories, each issue will include regular features such as First Ventures, Mapping how-to's and Tech Tips.

Benefits to the bi-monthly schedule are to provide greater time and space to develop each theme and to be able to coordinate more with *The IRE Journal*.

Our plan includes design improvements and to make available to subscribers an online, searchable archive of past articles. Subscribers will also receive news flashes via e-mail on developments in the world of data and hot new CAR stories available online.

MaryJo Sylwester can be reached by e-mail at maryjo@nicar.org

TRAINING

Census, maps and more

IRE and NICAR has a calendar full of opportunities this year, including a spring mapping bootcamp and CAR training sessions in Denmark and London.

We're also planning to schedule several more Census workshops in various locations of the country. Stay tuned for dates and places.

The mapping bootcamp, recently added to the schedule, will have a special emphasis on mapping Census data. It will be held May 4-6, in Columbia, Mo., at the Missouri School of Journalism.

Sessions by Andy Lehren of Dateline NBC and Jennifer LaFleur of *The St. Louis Post-Dispatch* will include thematic mapping, overlaying, spatial analysis and address geocoding. Participants are asked to have a basic understanding of database managers.

More information, including registration forms, is available at *www.ire.org/training/may4map.html*.

Also in May, IRE and NICAR will sponsor the Advanced Bootcamp on Statistics at the University of North Carolina-Chapel. See page twenty-two for more about this seminar.

In June, the IRE National Conference in Chicago will include a full-day of panels dedicated to computer-assisted reporting topics. The conference runs June 14-17. More information is available at *www.ire.org/chicago/index.html*.

And make sure to mark your calendars for the NICAR National Conference in Philadelphia, Oct. 11-14.

For European journalists, IRE and NICAR will offer CAR training at two conferences.

In April, IRE and other European journalism organizations will hold the Global Journalism Conference in Copenhagen. The conference, April 26-29, will include hands-on CAR training and panels on the latest reporting issues. More information is available at *www.ire.org/training/denmark/2000conf/*

IRE and NICAR will also offer CAR training at NetMedia, a journalism conference July 4-6 in London.

Finally, the usual lineup of NICAR Bootcamps are on schedule in Columbia, Mo. Seats are still available for the May, July and August sessions. More information is available at *www.ire.org/training/bootcamps.html*

Florida's dead heat

By Stephen K. Doig
Arizona State University

I left *The Miami Herald* for Arizona State University more than four years ago. But in November, for the first time, I wished I were back working in *The Herald* newsroom.

The reason was the dead-heat Florida presidential election. This political nightmare was a journalist's dream, and I knew that computer-assisted reporting would play a big role in the coverage. As the media frenzy began to build, I sent a "gee, if there's anything I can do to help, just yell" e-mail to *Herald* city editor (and IRE chairman) Judy Miller.

My eagerness to have a piece of this story must have been evident. A few days later, Judy called. We talked about different angles *The Herald* troops were pursuing, including an analysis underway by CAR specialist Geoff Dougherty indicating that hundreds of felons had voted in violation of state law. But in light of the evident confusion in so many Florida polling places, the burning problem at the moment was "discerning the mind of the voter" – particularly the clouded minds of the 185,000 or so Floridians who had gone to the polls but had failed for whatever reason to cast a valid ballot.

I did have an idea, albeit unformed, about how to attack the question. I had already been intrigued by how a smart statistician at Carnegie Mellon had demonstrated convincingly how vanishingly small was the probability that so many voters in Palm Beach County, home of the infamous "butterfly" ballot, would deliberately vote for Pat Buchanan.

Off the cuff

That analysis was based on summaries of the vote from each of Florida's 67 counties. Well, I figured, if we wanted to know what those failed voters intended, we should get closer to them. Instead of looking at county level results, I suggested off-handedly to Judy, we should gather precinct-level returns to see what they tell us. I didn't have a very clear idea of what I might do with such data, but it somehow seemed useful at the time.

It never occurred to me how staggeringly difficult it might be for *The Herald* to canvass all 67 counties, which were using all manner of voting methods and at least one of which was still using paper ballots. However, one thing *Herald* editors excel at is getting bright young reporters to work really hard when yet another "Story of the Century" breaks in South Florida. And so it happened with this one.

Two reporters, Anabelle de Gale and Lila Aruza, had worked long hours on the phones wheedling elections departments across the state, pulling together what I so casually asked for. Sometimes they got it as a nice, neat e-mail attachment. More often, it came as a fax from technologically backward – or deliberately obstructive – officials that had to be tediously retyped into a spreadsheet.

Making a table

As they completed each county, Anabelle and Lila e-mailed me the spreadsheets and I stitched those into one master Excel table that would cover the state's 5,885 precincts. For each precinct, the table showed the total vote for Bush, Gore, the combined other presidential candidates, and the total number of people who had cast ballots.

By the middle of the week after Thanksgiving, the pressure had shifted to me and they wanted a story for that weekend. And I didn't have anything to say.

My favorite [reaction] came from the Bush spokesman who dismissed it all as 'voodoo statistics'.

My initial plan had been to merge the results table with voter demographics from other sources, in the hope that we could make some definitive statements about who had been most affected by the ballot problems. But the demographic merge turned out to have huge problems caused by the varying ways each county named its precincts. It was a story that would be done later by Geoff, but the necessary cleanup put it out of reach

Continued on page four

The story discussed, "If The Vote Were Flawless... Gore Would Have Had The Edge In Glitch-Free Florida Balloting, Based On A Herald Analysis" by Anabelle de Gale, Lila Aruza and Curtis Morgan, published on December 3, 2000 in *The Miami Herald*, is available online at: *http://www.miami.com/herald/special/news/flacount/docs/104268.htm.*

More of *The Miami Herald's* reporting series on Florida Count can be found at: *http://www.miami.com/herald/special/news/flacount/index.htm*

The information about elections and voting in the U.S., including the official presidential general election results in 2000 can be found at the Web site of Federal Election Commission, *http://www.fec.gov/elections.html.*

IRE's Campaign Finance Information Center offers services for reporting on elections and money flow.

Visit the CFIC Web site, *http://www.campaignfinance.org*, to find state-by-state campaign finance data, tips, story ideas and training opportunities. You can also check out the Tracker, CFIC's newsletter at *http://www.campaignfinance.org/tracker*.

The IRE Resource Center has electronically compiled more than 1,200 campaign finance stories. To search the CFIC story library go to *http://www.campaignfinance.org/searchstories.html*.

Continued from page three:

Dead heat

for that weekend. For a while, all I could envision was some sort of squishy look at such trivialities as the number of uncounted votes in each precinct.

Eureka!

Finally, inspiration struck. Years of working with both elections and census data taught me that people behave pretty much like their neighbors do. A precinct, almost by definition, is a neighborhood. So it dawned on me to take the voting behavior of those who had successfully voted in each precinct, and just apply it proportionally to the uncounted votes. Thus, if 60 percent of the vote in a given precinct went to Bush, then I would give Bush 60 percent of the uncounted votes, and so on.

My analysis, and the ultimate headline, had focused on what would have happened if the vote had been flawless.

At that point, the analysis became a straightforward spreadsheet exercise, simple enough that I'm going to have my students duplicate it this semester. I did the formulas, copied them down the columns, hit the recalculation button, and saw the lead: Gore would have won by about 24,000 votes, instead of losing by about 500, if all those uncounted ballots had been cast correctly.

Reporter Curtis Morgan began bouncing our results off different sources for reaction. (My favorite came from the Bush spokesman who dismissed it all as "voodoo statistics"; I think I'm gonna have a T-shirt made with that slogan.) Although we were still refining the findings, Curtis managed to craft a nicely readable account with these initial interviews.

He also came up with the clever analogy of this analysis to a public opinion survey. Normally, a small sample is used to gauge the opinion of a large group. In this case, I was using the demonstrated behavior of the large group to project the behavior of a small subset.

My analysis, and the ultimate headline, had focused on what would have happened if the vote had been flawless. Some critics phoned by Curtis had retorted that a perfect election is impossible, and that therefore my analysis was meaningless. Others argued strenuously about the so-called "protest vote" of those who go to the polls but deliberately don't vote in a race because they don't like any candidate.

Another question

But *Herald* assistant managing editor Mark Seibel then raised an interesting question that hadn't occurred to me: How many uncounted votes could there be and still leave Gore ahead? Mark's question made me realize that this was a problem of limits. At one extreme was the actual situation: 185,000 uncounted votes and Bush winning by about 500. At the other end was my theoretical result: All votes counted and Gore ahead by about 24,000.

Even the most rabidly partisan Gore supporters would have to acknowledge that there's no way all 185,000 wanted to vote. But only the most rabidly partisan Bush supporters could pretend to believe that virtually none of the 185,000 wanted to vote. And the answer to Mark's question, I calculated, was that even if more than 90 percent of the uncounted ballots weren't counted, Gore still would win.

The story, which ran Sunday, Dec. 3, caused more fuss than any story in my career. I received more than 200 e-mails. Many were from those rabid Bush partisans, frothing at how I obviously was a hack political operative who had subverted the Constitution. A few were from Gore partisans who wanted me to give the data to his legal team; I declined. Curtis, whose e-mail address ran with the story, got more than 800 e-mails along the same lines. We also were deluged with requests for interviews, talk radio appearances, and such.

It was another week, when the Supreme Court stepped in and suspended the recount, before my 15 minutes of fame finally ran out. By then I was plenty sick of the attention; journalists can be a real pain, you know. But I sure had enjoyed the chance to play journalist again myself.

Steve Doig can be reached by e-mail at steve.doig@asu.edu

Four-day success

By Nancy Cook Lauer
The Tallahassee Democrat

Here's proof that you don't have to be a big newspaper with a lot of expensive database software and unlimited staff time to come up with a compelling story. From inspiration to publication, this CAR project took just four days - during which the reporter also churned out front-page dailies from Ground Zero of the 2000 post-election imbroglio.

As a newly promoted bureau chief at a small (just over 50,000 daily circulation) Knight Ridder newspaper in Tallahassee, Fla., I was up to my eyeballs in Florida's hanging chads and other controversies when a call came out of the blue. A civic activist I had known at a previous job five years ago and 400 miles away had driven to the state capital to get some information from the state Division of Elections. She saw my byline and called to tell me she'd picked up an interesting report.

To me, the primary challenge of a CAR project is balancing that desire to fine-tune, fine-tune, fine-tune, with the urgency of breaking news.

Because the information wasn't available on the state Division of Elections' Web site, I had a copy of the report faxed over. It was a compilation of the presidential votes cast and counted in each of the state's 67 counties.

Punch cards

Like the rest of the world, we'd been hearing that voters in counties with antiquated punch-card voting systems had fewer of their votes counted, compared to those using optical scanners. We'd also been hearing complaints from minority groups that poorer counties had the fewest votes counted because they had the bulk of the old equipment.

I used Microsoft Excel 97 (for PCs) and compiled a spreadsheet listing the state's 67 counties in one column and the percent of votes not counted in another. Generally, I prefer using raw numbers, not computed ones, in spreadsheets, but in this case, the percentages were all I needed and I was in a hurry to make a Sunday story out of information I'd just discovered on Thursday.

Because I was in a hurry and on the cheap, I went to the University of Florida's Bureau of Economic and Business Research site (*http://www.bebr.ufl.edu/*) and printed out a free report on population and income. Since it was a PDF, I had to key the per-capita income data back in instead of just downloading it.

I had to do the same with the information about which counties used what voter system. That's available on the Division of Election's Web site (*http://election.dos.state.fl.us/votemeth/cvs.shtml*), but not in any downloadable form.

Once the data were input and double-checked, there were several ways I could have approached the analysis. I could have done more intensive number-crunching on it such as Chi-square, but in the time frame allotted, I thought just ranking the counties by income and slicing them in half - high income vs low income - would offer a new perspective and serve the purposes of the article.

The resulting story "Incomes, discarded votes may be linked" (*http://web.tallahasseedemocrat.com/content/tallahassee/2000/12/03/local/1203.loc.votes.htm*) showed that voters in Florida's poorer counties were more than twice as likely as those in more affluent ones to have their votes for president disregarded. That held true regardless of whether they voted on optically scanned ballots or punch card ones.

The story inspired debate about the get-out-the-vote campaigns that had been launched by the National Association for the Advancement of Colored People, unions and Democrats, all of whom tend to draw lower-income voters. Advocates used the article to talk about the need for voter education and the equitability of having the state subsidize new voting machines. Even Gov. Jeb Bush found the article of use and had a copy faxed

Continued on page eleven

The Tallahassee Democrat's "Incomes, discarded votes may be linked" is available online at *http://web.tallahassee democrat.com/content/tallahassee/2000/12/03/local/1203.loc.votes.htm*

"Bush wins battle for legal funds" is also available online at *http://web.tallahassee democrat.com/content/tallahassee/2000/12/18/local/1218.loc.contribs.htm*

These stories will soon be available from the IRE Resource Center. Search the story database at *www.ire.org/resourcecenter* for order numbers. Call (573) 882-3364 to order these and other investigative stories.

Election machine

Neil Reisner
Miami Daily Business Review

Now that Chad-Fest 2000, otherwise known as the presidential election is over – I suppose I could call it a Chad Row, but awful puns like that are just so tacky – it might be a good time to come back to earth and think about the kinds of elections we normally cover.

Small stuff: town council, school board, county commissioner, maybe governor.

And lots of computer savvy reporters figured out a long time ago that a spreadsheet is an election reporter's best friend, ideal for tabulating votes, percentages and winners and able to produce neat little graphs lickety-split.

But as soon as results are in, the questions start coming:

What happened? How'd she win? or How'd the Black community/Hispanic neighborhoods/the rich/the poor/home owners/condo residents/gays/swing precincts vote?

Usually that means making a quick checklist of key precincts, scanning the spreadsheet and tweaking out whatever patterns may be visible. It's not easy and there's a better way. A map of the election will often let the patterns jump out at you and your colleagues whose eyes may glaze over when they look at a table of numbers. It's not hard, but to make it happen requires some front-end work before the election. I call it the "Election Machine."

You'll need:

- Excel, Access and ArcView or some other spreadsheet, database manager and mapping program.
- Precinct-by-precinct voter registration data, in the spreadsheet.
- Tables of key voting precincts, in the database manager.
- A map layer showing election precincts.
- Election results in the spreadsheet.

Here's a step-by-step guide to how it all fits together using Excel, Access and ArcView. Looking at real examples will make it a little easier to follow; if you want to try if for yourself, download a PowerPoint demonstration and sample data in Excel, Access and ArcView from the NICAR Web site: *www.nicar.org/electmach/index.html.*

Step 1: Voter Registration

Build a one-page workbook showing the number of Democrats, Republicans, etc. in each precinct or election district, as well as the number of blacks, whites, Hispanics and any other ethnic or demographic information you can obtain. The best source for this usually is the county supervisor of elections, who should keep this information on each precinct; other sources might be local political consultants or university political scientists or demographers. Figure the percentages of each political or demographic group: Democrat, Republican, black, white, etc.

Step 2: Key Precincts

Create a database and build a table identifying key precincts, bellwether precincts that are predominantly Democratic, Republican, Black, white and the like.

Step 3: Election Results

Build a spreadsheet workbook with one page for overall returns, precinct-by-precinct and a separate page for each election that you want to analyze. Why? If you're lucky, you can get results in some electronic format, through e-mail, a download or on a floppy disk.

Ask the election folks for a dummy file before – the same layout you'll see on Election Day, but without numbers – and build a dummy page mimicking that layout. Then it becomes a (reasonably) simple matter to paste the results onto the page. Try this in advance to make sure it works. Worst-comes-to-worst, build the page in advance and type results into it.

If you're dealing with multiple elections – several municipalities or commission seats, etc. – copy results for each election in which you're interested to separate pages, one page per election. Paste the results from the results page as a live link by clicking on "Edit," "Paste Special," "Paste Link." When you do it this way, the results of each individual election will flow automatically from the results page to the election pages.

Build formulas, calculating what you want to know about the election: the turnout, the winner, etc.

Continued on page seven

Tipsheets for local elections:

Tipsheet # 923. Francis Gilpin of the *Tampa Tribune* examined how CAR can help discover the power of local lobbyists in elections. (1999, Boston).

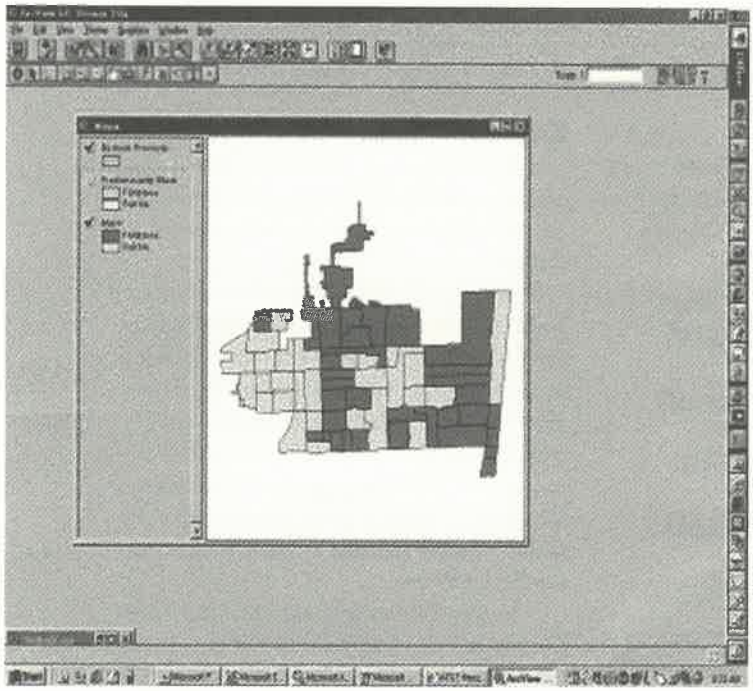
Tipsheet # 924. David Washburn offered tips on local reporting including past stories on commissioner salaries and elections. (1999, Boston).

Tipsheet # 478. This tipsheet includes a list of election story ideas as well as tips on setting up a contributions database.

Continued from page six: **Machine**

How Come...

• ...it's important to keep registration numbers in a separate workbook? Registration data will change election-to-election and you'll sometimes have to re-calculate percentages. Keeping it separate lets you build a registration template, updating it for each election as the numbers change



• ...I should keep overall election results on one page and each election on others? Because it lets you paste/type all the results on one page instead of having to do it on multiple pages. Because it lets you type all results on one page.

• ...I should keep key precincts in a database? Because key precincts are not likely to change and you can use the same base data for each election.

Step 4: Name Spreadsheet Election Pages

On each individual election's page, select the range containing the results, including headings. Give it a name – City Council, for example – by clicking "Insert, Name," Define" and "Add." Do the same for the voter registration page. You'll see that the name of the range will appear in a small window left of Excel's formula bar.

Open Access and click "File," Get External Data," and "Link Tables." Select the election results workbook. When the Wizard opens, click on "Show Named Ranges." You'll see that the ranges you just named in Excel show up as items you can link to Access. Follow the wizard's instructions and link each Excel range to Access as a table.

Do the same in the registration workbook: Name the range containing the demographic percentages in which you're interested and link it to Access as a table.

Step 5: Build and Run Queries

Build queries to look at winners, losers, ethnic votes or anything else of interest. Link to swing precincts winner, Black vote or anything else of interest. Link precinct-by-precinct registration numbers to precinct-by-precinct results table or swing precinct table to determine voting patterns. You can build the queries in advance, linking to your dummy data. That makes things faster on deadline.

Step 6: Export Queries

Export the results of each relevant query in dBase IV (*.dbf) format by clicking on "File," and "Save As/Export." Choose "To an External File or Database" and follow the wizard. This will let you import the results into ArcView.

Step 7: Build Maps

Open ArcView, where you've already built a precinct map. You may be able to get a precinct map from your local election supervisor, city or county planning department, the state legislative staff responsible for redistricting, political consultant or political scientist may have one.

Join the *.dbf tables you exported from Access to the map.

Continued on page nine

Other local election stories:

Rose Ciotta of the *Buffalo News* (Buffalo, N.Y.) examined political fundraising for a 1995 county executive race. Story # 12358.

In 1994, the *Times* (Munster, Ind.) investigated absentee ballot vote fraud in the spring primary election. Story # 10944.

"GOP-petition signatures questioned," by Greg Moran of the *Post-Star* (Glens Falls, N.Y.), Aug. 1987, found forged signatures on the petitions to get local GOP candidates on the ballot. Story # 5119.

Copies of these stories can be ordered from the IRE Resource Center by calling (573) 882-3364. Search the story database at *www.ire.org/resourcecenter*.

CAR in Election 2000

The New York Times and the Scripps Howard post-election analysis discussed in Ron Nixon's story on page eight will be available soon at the IRE Resource Center. The contact number to the Resource Center is (573) 882-3364

By Ron Nixon
IRE and NICAR

As the 2000 Presidential election dragged on due to lawsuits filed over accusations of missed and uncounted votes, Thomas Hargrove, a reporter for Scripps Howard, wondered if the 2000 Florida fiasco was an isolated problem.

To test his theory, Hargrove went to Election Data Services, a Washington group that keeps election results, for data from the 1996 election.

Pulling the data into SPSS, a statistical software package, Hargrove ran queries to see if there were indeed undercounts in the contest between Bob Dole and Bill Clinton.

The results were startling.

Computer-assisted reporting played an important role in coverage of the 2000 presidential election.

"Our investigation showed that in the Dole-Clinton race, millions of votes were not counted," he said. "In one West Virginia County, the county clerk had just found 1,000 votes that Bob Dole should have gotten.

"It wasn't enough to change the election because Clinton had such a lead over Dole," Hargrove said.

The News Service also found other problems: official registration lists in the United States have become so faulty that scores of counties claim to have more voters than actual adult population.

Sloppy bookkeeping led 190 counties and the state of Maine to appear to be more than 100 percent registered in the 1996 presidential election, according to the Scripps Howard News Service study of election records in 3,189 counties and voting districts.

"Our studies did show that there were numerous problems in voting across the country and not just Florida," Hargrove said.

Computer-assisted reporting played an important role in coverage of the 2000 presidential election.

Hundreds of newspapers and other media across the county found problems in the election system that went beyond Florida, which had gathered most of the attention in the presidential race. Among the papers and news organizations were *The New York Times*, *The Wall Street Journal*, *The Miami Herald*, *The Palm Beach Post*, *The Chicago Tribune*, and *The Tallahassee Democrat*, among others.

Key findings

• The long-held belief that every vote counts is a myth. Media organizations found that in each election, thousands of votes are thrown out or not counted.

• The 3,067 individual counties in the United States have few national standards for determining how votes get counted. The job of counting the 100 million or more votes often falls on poorly trained and understaffed county clerk offices with few resources. A *Wall Street Journal* investigation found that county governments typically spend about 3 percent of their budgets on the clerk offices.

• Voting failures often fall on those who can least afford it: Minority and low-income communities. Investigations by *The New York Times* and several papers in Florida show that voting systems in these communities tend to be older and less reliable than those in high-income communities.

• Registration lists were often flawed. News organizations found that many people were turned away from polls because their names weren't on the list, and a *Miami Herald* investigation found that hundreds of felons voted in the presidential election.

• In Florida, poor ballot designs may have given thousands of votes to the wrong candidate. Another *Miami Herald* investigation found that a third party candidate's gain in votes in a heavy Democratic county was a statistical anomaly.

Faulty machines

Most of the problems with under-votes were attributed to faulty voting machines, an investigation by a media organization found.

One of the first newspapers to look at the

Continued on page nine

Continued from page seven: # Machine

The example here joins a "Winner" table to the precinct map. It shows that Rubble didn't get beyond his political base, while Flintstone took his base in Central Bedrock and more. [see graphic]

Summing up

Using a spreadsheet, database manager and mapping software you can quickly create a picture of what happened in an election. Just gather the basic data that you'll need in advance and build spreadsheets, tables and queries.

Build dummy spreadsheets for the election results you won't get until Election Day. When results come in, you can paste them or type them on a single spreadsheet page and everything will flow from Excel to Access and into pre-constructed queries. Export the results, join them to a precinct map and you've got it. This may seem very complicated, but once you get the hang of it, it's not.

Some Other Possibilities

- A map that lets you quickly see how an election was won or lost.
- Refine the analysis by examining voting blocs.
- Examine candidates' margins to look at the election more closely.
- Compare this election to past elections.

Neil Reisner can be reached by e-mail at Nreisner@floridabiz.com

Continued from page eight: # CAR in Election

problem was *The Tallahassee Democrat*. [see full story on page five]

"We kept hearing from community groups that something was wrong with the machines," said Nancy Cook Lauer, now Capital Bureau Chief for *The Tallahassee Democrat*. "We wanted to check it out, to see if there was indeed a problem or if this was just partisan bickering."

The New York Times analysis showed that there were problems with voting machines and uncounted votes. Blacks were more likely to have their votes discarded because they often lived in counties that used more error-prone punch cards, the paper found. Sixty-four percent of black voters lived in counties that used punch cards as opposed to 56 percent of white voters, the *Times* found.

An investigation by *The Herald* found that Reform Party candidate Pat Buchanan's 3,400 votes in a heavily Democratic county was a statistical anomaly. [see full story on page one]

Even after the election was decided, news organizations stayed on the story using data analysis skills to find answers. *The Palm Beach Post* conducted a ballot-by-ballot review of discarded over-votes in the county.

The results, published in early March, found the confusion over the butterfly ballots cost Gore about 6,600 votes. That would have been more than 10 times what he needed to overcome Bush's slim lead in Florida and win the presidency.

A team of 10 *Post* reporters reviewed the ballots Jan. 17-29 at the Supervisor of Elections Office in West Palm Beach. The county charged $125 per hour, which was split among the paper, the Republican Party and two other groups also doing reviews.

After collecting data on handwritten tally sheets, *Post* reporters entered their findings into a computer database, recording the names of the candidates selected.

To verify the data and eliminate entry errors, a staff writer compared *The Post's* database to the official precinct-by-precinct number of over-votes recorded after the county's 10-day hand recount in November.

Those numbers were off by between 17 and 117 ballots in 14 precincts. The county wouldn't allow another review of the ballots until after April 2 (because of municipal elections), so the paper supplemented its database in 13 of those 14 precincts with data from a computer record of the ballots produced Nov. 11 after the third machine recount in Palm Beach County.

The data came from University of California-Irvine researcher Anthony Salvanto, who is studying the election results. Those numbers matched the county's hand recount results.

Ron Nixon can be reached by e-mail at ron@nicar.org

WTVF-TV, Nashville, reported on dead voters. To read the story go to this web site:

http://www.newschannel5.com/news/investigates/vote.htm

To read the *Palm Beach Post's* coverage, go here:

http://www.gopbi.com/partners/pbpost/news/election2000.html

Software:
Microsoft Access

Data:
Hospital billing records: Illinois Health Care Cost Containment Council (IHCCCC).

WMAQ-TV's tape will soon be available in the IRE Resource Center. Search the story database at *www.ire.org/resourcecenter* to find the order number. To order this and other stories, call (573) 882-3364.

A sample of the data used can be downloaded from the NICAR Web site, *www.nicar.org/techtips.html*

Hospital Infections

By Michele Rubenstein
WMAQ-TV, Chicago

So I told my news director, "Send me to NICAR Bootcamp and it will be worth the money!" I promised to use what I learned to uncover stories we otherwise would not find. But when I returned, and it was time to produce, I thought I'll never be able to do this!

My first project, I decided, didn't have to be groundbreaking. It had to be simple. We bought an easy-to-use database of hospital billing records to find out how many patients had been diagnosed with infections while being treated for other illnesses. Wouldn't you know, while working on this story we uncovered information that had never been reported. We found a potential outbreak of a deadly infection.

The story

Necrotizing Fasciitis, known as flesh eating bacteria, deforms or even kills its victims. It can be highly contagious and, in its most deadly form, can kill an otherwise healthy person within 24 hours. Because this infection can be communicable, it is on the Illinois Department of Public Health's list of reportable diseases.

State health officials say the infection is rare, showing up in only a handful of people. But we discovered more than 500 cases of Necrotizing Fasciitis.

"The number is large. It's much larger than we would expect," said Dr. John Lumpkin, director of the Illinois Department of Public Health. As a result of our story, he launched an investigation to find out if there has been a silent outbreak of the infection and if hospitals are failing to report a reportable disease.

Ironically, we found this story while working on a different series involving hospital-acquired infections. We obtained a database of hospital billing records to track the number of infections reported at each hospital. Before narrowing the search, we spent time playing with the data. That's when the story of Necrotizing Fasciitis revealed itself.

About the data

We obtained our database from a little known state agency called the Illinois Health Care Cost Containment Council (IHCCCC). This agency tracks hospital billing costs so consumers can find out if they are being correctly charged for their care. So clearly, this database can be used to find multiple stories.

The records we were given did not include patients' names, but did list age, gender, county they were from, primary diagnosis, secondary diagnosis, procedures and more. It was easy working with the records. The only problem was interpreting the diagnoses and procedures, which were coded. For this, we needed a second database.

State health officials say the infection is rare, showing up in only a handful of people. But we discovered more than 500 cases of Necrotizing Fasciitis.

When doctors bill for medical treatment they use ICD-9CM codes. Each diagnosis and treatment is assigned a code number. Therefore, you need an ICD-9CM database. I was told we could purchase an advanced computerized version of this for about $500. I did not have the time to wait or money to spend to get this database. Fortunately, I found someone who e-mailed a simplified and complete version to me for free. For more information about the codes, you can call the American Hospital Association, which has ICD-9CM experts.

A word of warning about working with ICD-9CM codes — there are inpatient codes and outpatient codes. There are different codes for diagnoses and procedures. You need to make sure you are working with the right set. Also, there can be multiple codes for a condition. For example, if you want to find out how many patients have been diagnosed with blood stream infections, you would have

Continued on page eleven

Continued from page ten: **Hospital**

to look up eight different ICD-9CM codes.

Obtaining data

This story can be quickly done once you have the data. But getting the records can be difficult. It took months to get our main database from the IHCCCC. Most of the time was spent on wasted phone calls and pointless correspondence with the agency's publicist. It took pushing them into setting up a meeting with the computer specialist to finally get our records.

The next problem was the cost. The agency wanted to charge us about $6,000 for the records. That's what they charge hospitals and other corporations. I obtained a copy of the fee structure and discovered educational institutions receive a discount. I was able to argue we should be given the data at a discounted price, since we were going to use it to educate our viewers. But, to get the cost under a thousand-dollars, I had to make one compromise. I had to give up getting the entire database. Instead the agency did a run for us pulling the records of patients who had an infection as a secondary diagnosis. For the story we were doing, that's all we needed.

Once you get your records and your ICD9-CM codes, you should be able to turn your story rather quickly. To find "victims," you can get help from Jury Verdict. They track lawsuits and will conduct special computer searches using key words such as "flesh eating" and "Necrotizing Fasciitis".

To do this story, you only need Access. It was relatively easy to produce and struck a chord with our viewers. We received numerous tips for follow-up stories and we are tracking the state health department's investigation into our findings.

So, I guess your first computer-assisted report can be simple and still be groundbreaking. I'm just glad mine is over.

Michele Rubenstein can be reached by e-mail at michele.rubenstein@nbc.com

---

Continued from page five: **Four-day**

over to him a month later, the morning of his testimony before the U.S. Commission on Civil Rights.

Post-election money

Buoyed by my success with the voting project, I undertook a bigger one - an analysis of post-election contributions to the Bush and Gore campaigns. (*http://web.tallahasseedemocrat.com/content/tallahassee/2000/12/18/local/1218.loc.contribs.htm*) I wanted to see which states gave the most money to each candidate during the recounts.

This project was inspired by a press release from the Gore campaign touting the number of contributors giving small donations. I went to the Gore recount Web site (now closed) and on Dec. 9 downloaded the 19-megabyte jpg file of the campaign's Form 8872: Political Organization Report of Contributions and Expenditures. It was free, but cumbersome because I then had to input information about each contribution, and the Gore campaign insisted it had no other format. Over a weekend, I input the state and amount for 1,258 contributors into an Excel spreadsheet.

Bush's information was much easier - I just downloaded a zipped version of his Excel file containing about 22,000 contributors from his Web site (*www.georgewbush.com*). After cleaning up the data, I imported both the Bush and Gore numbers into a Microsoft Access 97 (for PCs) database to make summary tables. I filtered everything out of the Bush database except the state and amount while importing it into Access.

This project lost a lot its luster halfway through - The U.S. Supreme Court ruled, Gore conceded and I wasn't sure anyone cared about the contributions anymore. I pushed forward anyway and had the story out just over a week after getting the first information. To me, the primary challenge of a CAR project is balancing that desire to fine-tune, fine-tune, fine-tune, with the urgency of breaking news.

Nancy Cook Lauer can be reached by e-mail at nlauer@taldem.com

---

There are more than 200 news articles reporting on hospital-related topics. Search the story database at *www.ire.org/resourcecenter*.

Story #12148: "When Cure Becomes Cause: Hospital Hazards" by Adam Berlaint and Elaine Porterfield for News Tribune on May 28, 1995.

Helpful Web sites related to the story:

Illinois Department of Public Health, *http://www.idph.state.il.us/*

American Hospital Association, *http://www.aha.org/index.asp*

METHAMPHETAMINE

Ammonia leaks

By Dawn Fallik
St. Louis Post-Dispatch

It's hard to escape a leaking ammonia tank. The smell of an unchanged litter box permeates every sense. Your eyes water and your nose burns. If you inhale too much for too long — and that can be seconds — your nose hairs fall out and the lining of your throat starts to blister.

In Missouri, and in many rural states, those leaks, once rare, now occur regularly at farmer's cooperatives and on farms. Although many releases are caught early, some result in emergency evacuations, injuries and occasionally, fatalities.

The rise of methamphetamine, a popular drug similar to speed that's easy to make and cheap to buy, use directly relates to the increase. The same liquid farmers use to add nitrogen to the soil is also used to make methamphetamine. Because farmers and agricultural cooperatives use such large tanks — at least 10,000 gallons — and the amount of ammonia needed is so small — only a few tablespoons, most ammonia thefts go undetected.

As methamphetamine use increases, thefts do too, and drug addicts, not being trained in hazardous material handling, leave valves open and cut lines, leaving the toxic gas to spew into the air.

In the past five years, there have been more than 200 ammonia releases in Missouri — more than any other single chemical. While researchers try and create a type of ammonia that cannot be used for meth, police officers struggle to handle hazardous materials they never encountered before, and agriculture officials try to stop the vandals, without much success.

Discovery

I stumbled upon the story while looking for something completely different. I had attended a hazardous material spills conference in St. Louis and discovered Missouri took part in a Hazardous Substances Emergency Events Surveillance program. I called the Department of Natural Resources to find the program coordinator — only to find a completely different database — a detailed description of all the events called into the state emergency response center.

I obtained both the HSEES (kept by the Department of Heath, not DNR) as well as the emergency events data. Both datasets came in Access format and neither agency charged me a cent.

As I browsed through the events database, I noticed there were many incidents at agriculture facilities, and they were marked as "intentional releases." In looking at the HSEES set, I found that several of the releases the DNR mentioned did not appear.

Investigating the data

Then I noticed a few of the incidents at the agriculture facilities also mentioned evacuations. In one incident, more than 250 people had to be evacuated at 4 a.m. from a town about 30 miles south of Kansas City. In another, a man was killed when a fire extinguisher, illegally filled with anhydrous ammonia, exploded in his lap while driving on an interstate in the middle of the day. A firefighter, a paramedic and a passer-by who stopped to help all suffered respiratory irritation and inhalation burns.

I also noticed several incidents at jails and in courthouses. It turned out that police officers, not knowing anything about hazardous materials, caused a few evacuations by placing illegal ammonia containers, which subsequently leaked, in empty jail cells. One police officer's office caught on fire when he left some water-reactive chemicals on top of a filing cabinet and the roof began to leak.

Witnessing data beginning

I ended up writing a mainbar on the increase in ammonia leaks, focusing on recent evacuations and injuries, as well as sidebars on toothless meth laws and the new challenges facing emergency workers who were dealing with hazardous materials, as well as violent drug addicts. I accompanied an officer on his rounds of local ammonia facilities and ended up getting a blast of ammonia fumes — while the officer and I were walking around the co-op, someone had broken into an outside tank.

Although environmental officials gave me the technical breakdown, police officers made the story — telling tales of finding meth labs under car seats, ammonia in dishwashing liq-

Continued on page thirteen

Software used: Microsoft Access

Data: Hazardous Substances Emergency Events Surveillance program and an emergency events database.
Source: Missouri Department of Health, Missouri Department of Natural Resources

Other stories on hazardous materials: These are available from the IRE Resource Center, (573) 882-3364.

Story # 8795. *The Los Angeles Times* conducts a database study of 68,000 hazardous materials incidents from around the United States, and finds the number of incidents and injuries rose between 1982 and 1991.

Story # 14184. *The Washington Monthly* reports that serious accidents involving the transport of hazardous materials on trucks and trains have become an almost daily occurrence.

Continued from page twelve: **Leaks**

uid containers and having the toxic liquid thrown in their faces as thieves escaped. Officers also explained how frustrating it was to catch people with all the makings of meth — but not being able to arrest them because none of the materials are illegal.

Results: additional training

Until last summer, the only agency training police officers in hazardous material handling was the Drug Enforcement Agency. But that training only occurred a few times a year and seating was scarce.

Then last summer, the Missouri Highway Patrol and the Department of Natural Resources teamed up to offer free classes on meth and the chemicals involved. At the same time, police officers began working with farmers and agriculture facilities to help stop the ammonia leaks. Simple precautions such as lighting and locks help some, but, after the story ran, I heard from a lot of farmers saying that simple precautions weren't doing much anymore.

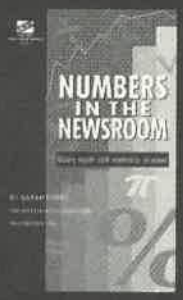
Although I used the HSEES and the emergency events database solely for ammonia releases, both include many more details of hazardous cleanup and response actions. Of course, both datasets had factual problems and dirty data, but they were good tips to what because an unusual angle on a now-common problem.

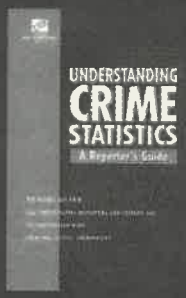
Dawn Fallik can be reached by e-mail at dfallik@post-dispatch.com

REQUIRED READING FOR YOUR NEWSROOM The IRE Beat Book Series

NUMBERS IN THE NEWSROOM:
Using math and statistics in news

THIS BOOK INCLUDES:
Writing about fractions, rates, percents and per capitas
Measuring change
Makin inflation adjustments
Understanding averages
Working with graphics
Doing the budget story
Questioning surveys and polls
... and much more



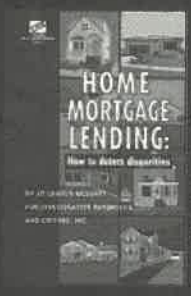


UNDERSTANDING CRIME STATICS
A Reporter's Guide

BOOK CHAPTERS:
Crime by the Numbers
Uniform Crime Reports
National Crime Victimization Survey
National Incident Based Reporting System
Other Major Statistical Sources
Writing the Crime Statistics Story
Database Analysis of Crime Statistics
Web Sites
Resources: Tipsheets and Stories

COVERING AVIATION SAFETY:
An Investigator's Guide

Learn to develop a newsroom crash plan, report from the scene of a crash, start an aviation beat, interpret aviation records, negotiate web-based data, and investigate planes and airlines on deadline.



HOME MORTGAGE LENDING:
How to detect disparities

Pulitzer Prize-winning journalist Jo Craven McGinty guides reporters through understanding and using Home Mortgage Disclosure Act data. Included are specific story ideas and lists of recent tipsheets and stories available through IRE.

ORDER NOW!
IRE MEMBERS: $15 NON-MEMBERS: $20 EACH

BY PHONE
Call 573-882-3364 with your VISA or MasterCard

BY MAIL
Send your check to IRE
138 Neff Annex
Missouri School of Journalism
Columbia, Mo. 65211

BY WEB
Visit our Web site at www.ire.org for online ordering or order form downloads

For more on reporting chemicals, check out these tipsheets found in IRE's Resource Center:

Tipsheet #1225, "Crashes, Explosions and Spills: Chemicals/Medical/ Science on Deadline," by Dawn Fallik of the *St. Louis Post-Dispatch*, provides a list of sources to have on-hand as well as a list of contacts and questions to ask when reporting on environmental issues.

Tipsheet # 987, "Environmental Perils Toxic Chemicals at Home and at Work," by Elizabeth Alex of KSHB-TV gives an outline of how to successfully investigate environmental issues.

Tipsheet #988 by Sam Roe of the *Toledo Blade* offers information on how to track victims of environmental hazards.

These tipsheets and others are available from the IRE Resource Center by calling (573) 882-3364.

Check out these tipsheets for ideas on how to deal with dirty data

Tipsheet #944, "Bad Data Does Not Make Good **Stories" provides information on how to deal with dirty data and campaign finances.**

Tipsheet #695 covers the basics of how to clean data. The handout focuses on dealing with string functions and dates as well as reading syntax descriptions.

Tipsheet #67 by Brant Houston discusses a technique for cleaning up inconsistent or dirty data.

To order these and other tipsheets, call the IRE Resource Center, (573) 882-3364.

Continued from page one:

Aftermath

moved because their information matched that of people listed as dead in the state's vital statistics register and the Social Security Death Index.

• Computerized maps, obtained from the Florida Legislature's Web site (*www.leg.state.fl.us/edr*)showing the location of each Florida precinct, along with demographic information and other geographic features.

• Data files from county election supervisors containing information on the voting patterns of each ballot cast in seven of Florida's largest counties.

• Files containing information on each ballot reviewed by the accounting firm hired by *The Herald* to inspect all of Florida's undervoted ballots. Those efforts began when reporters in Palm Beach County wanted to follow up on complaints that elderly voters had been confused by the ballot design. Many said they had accidentally voted for Pat Buchanan instead of their preferred candidate, Al Gore.

I used SAS statistical software to analyze the precinct returns in Palm Beach. Almost immediately, it was clear something was amiss. Voters in Palm Beach had turned out for Buchanan in numbers far higher than other counties in the state - and in much larger numbers than they had in previous elections and the 2000 primary.

Regression analysis

How to quantify that? I turned to a statistical technique called regression analysis, which allowed me to use the Gore/Buchanan voting pattern across Florida to predict how many votes Buchanan should have received. The next day's *Herald* reported that Buchanan received 2,000 more votes in Palm Beach than one would have expected by looking at the Gore totals.

From there, colleagues Tim Henderson, Jason Grotto and I turned to the question of illegal voters. We encountered difficulties immediately; Florida law says county election supervisors are not allowed to release databases that show identifying information about voters. The state Division of Elections can do so, but that file wouldn't be available for months.

Though a variety of means, we got our hands on voter files for several of the largest Florida counties. We matched the names in those files against our felons file, the death information, and against voter files obtained from other states around the country. Eventually we found thousands of felons had voted in the election, despite a Florida law barring them from doing so.

Our efforts to identify dead voters and those who had voted twice - either in two Florida counties or in Florida and another state - were slowed by a number of data problems. Some states didn't have information on who had voted, so our matches showed only that someone had voted in Florida while being registered in another state. None of the files identified voters by Social Security number, so our matches by name and date of birth had to be verified through time-consuming interviews and online research.

And our attempt to find non-U.S. citizens who had voted was stymied by another data flaw: After identifying themselves in court papers or on jail logs as non-citizens, many people went through naturalization proceedings and were U.S. citizens by the time they voted.

Interviews

While reporters spent time in the field tracking down potential illegal voters, others were interviewing legal voters who said their attempts at casting a ballot had been foiled - by poor ballot design, mechanical problems with voting machines and uncooperative pollworkers.

Many of those complaints came from heavily African-American precincts in urban Florida counties.

To get a handle on how widespread those problems were, we married information on the discarded ballot rate in each precinct in Florida to racial data from the state's voter registration files.

The first story to come out of that work was published the week after the election and concluded there was an unusually high rate of overvoting in Miami-Dade's minority precincts. Later, *The Herald* published stories documenting the same problems in Jacksonville.But was it enough to say that ballots belonging to black voters were discarded frequently? We decided to find out why those ballots were discarded.

Continued on page fifteen

Continued from page fourteen: **Aftermath**

To do so, I used the voter registration files to produce precinct summaries showing the average age, racial makeup and party affiliation. *Herald* news clerks called each county election supervisor to find out whether the voters there used punchcard systems or optical scan machines, and whether the ballot in each county had listed all the presidential candidates in a single column, or spread them across two columns.

Finally, I used Arcview mapping software to identify the nearest school to each precinct, and then used that school's free- and reduced-lunch population as an indicator of poverty in the precinct.

Another regression analysis led to some newsworthy findings: While voters in African-American precincts saw their ballots discarded frequently, race wasn't the entire story. Poverty, age of the voter, voting system and ballot design were closely correlated with the rate of ballot spoilage.

While punch-card voting systems were blasted in the weeks after the election, *The Herald* analysis showed optical scan systems that tabulated ballots at a central location were actually worse. On the other hand, optical systems that scanned ballots in each precinct and rejected those with overvotes and undervotes had a much lower rate of ballot spoilage.

Lessons learned

The Herald continues to analyze data from the election, and will soon publish the results of the statewide ballot review.

Meanwhile, the past three months have provided some lessons for us. Chief among them: It's difficult to coordinate efforts among CAR specialists if they are using different software and there is no central data-storage facility.

At *The Herald*, we've got a wide array of tools to use on data: SAS, SQLServer, Access, Foxpro, and others. But on projects like this one, which require the work of several database people, it's best to pick one piece of analysis software and make sure everyone uses it. That way the time spent converting files back and forth is minimized, and each reporter can follow the others work, checking and making suggestions where necessary.

Before the election, *The Herald* had discussed purchasing a large server that could house Census files and other CAR data. But we didn't have the server on Nov. 8, and we quickly found ourselves in a jam. Some files were so large (the state voter file exceeds 1 gigabyte) that there's no efficient way to move them from one computer to the other – which means that the work one reporter has put in importing and cleaning files is duplicated in order to allow another reporter to work with the same data. So get a big server.

Dirty data

We learned – or were reminded once again – that all data is dirty. Some of the people identified as felons in the corrections database weren't. Some of the people identified as Nov. 7 voters weren't. Many of the people listed in Social Security records as dead are not. As they will tell you, they're alive, well, and happily enjoying their retirement benefits in Florida.

Fortunately, all of these problems cropped up before we printed stories based on faulty data. We found the data flaws by making sure that we verified a healthy sample of the data by checking paper records and interviewing each person identified in the story. It's tempting to skip these time-consuming steps when the competition is intense, as it was after the election. But it paid off every time.

Finally, establishing a relationship with academics who have social science expertise is key. When I finished the regression analysis that looked at Buchanan and Gore votes in Palm Beach County, it was about 7:30 p.m. and editors were demanding the final numbers. But I hadn't yet talked about the statistics behind the analysis with anyone who knew enough to point out problems.

Fortunately there's a professor in the Florida university system who's good at that sort of thing. And because I've worked with her on a number of stories like this, I had her home phone number and knew she'd be willing to quickly determine whether our analysis was on solid statistical ground. I called her, talked for 10 minutes and punched the story.

Geoff Dougherty, who is starting a new job at the Chicago Tribune in April, can be reached by e-mail at geoffdougherty@yahoo.com

Sources for disregarded votes data:

University of Florida Bureau of Economic and Business Reasearch site, *http://www.bebr.ufl.edu/*

Division of Election Web site, *http://election.dos.state.fl.us/votemeth/cvs.shtml*

George W. Bush campaign contributions, *www.georgewbush.com*

10 Census stories

To find a 1990 Census tract for a particular address:

Punch in the address at *http://tier2.census.gov/ctsl/ctsl.htm* **or** *http://factfinder.census.gov/java_prod/dads.ui.map.ref.RefMapPage*

Libraries that store Census microfilm: *www.census.gov:80/prod/2/gen/cff/cff-9702.pdf*

To look up a county or town's land area and density in 1990, *www.census.gov/population/www/censusdata/density.html*

1990 Undercount and Overcount: *http://tier2.census.gov/pl94171/pl94index.htm*

IRS County-to-County Migration Data: *http://ftp.fedworld.gov/pub/irs-soi/prodserv.pdf*

By Paul Overberg
USA Today

Editor's note: This is a tipsheet prepared for the 2000 National CAR Conference in Lexington, Ky.

To cover Census 2000, it helps to have zigabytes of storage, light-speed processors and a barn full of computer memory. But it's not a requirement. The more planning you do, the better the stories you can do with limited resources. Here are 10 ideas that won't take anything fancier in the digital realm than a spreadsheet (free versions on the Internet, if it comes to that) and a Web browser. It presumes that you have a way to extract Census 2000 data you need, either by yourself or with help from another journalist or a state data center, local university, planner, etc.

Here are 10 ideas that won't take anything fancier in the digital realm than a spreadsheet and a Web browser.

Materials you need: Big paper wall map of census tracts in your county. Calculator. Spreadsheet. Phone numbers of local librarian, planner, suitable source lubrication (chocolate, home-baked goods, cup of coffee, complimentary newspaper subscription, etc.).

Roots

Find the oldest house in your town. Find out what census tract it sits in. (For its 1990 tract, punch in the address at *http://tier2.census.gov/ctsl/ctsl.htm or browse down at http://factfinder.census.gov/java_prod/dads.ui.map.ref.RefMapPage.*) Get data from that tract as far back as you can – possibly 1940. Then go to the nearest library that stores census returns on microfilm (See *http://www.census.gov:80/prod/2/gen/cff/cff-9702.pdf.*) The most recent census open to the public is from 1920. If the home was around then, pull the individual returns from that street, and go back to previous censuses as far as possible – 1910, 1900. ... You'll probably find surnames that have become street names, and descendants of those people probably still live in your community. Find them. A historical group may have done some of this – check. Use the data, the facts and the contacts they represent to write a story about the changes to that neighborhood.

A Road Runs Through It

Pick a major street that runs a long way through your community. Chances are it passes very different kinds of neighborhoods. Make a list of the census tracts it passes. Pull together simple 1990 Census data about them. Do the same when the first 2000 Census data comes out this summer. Write a story about how the road links places that are different and separated by only a few miles.

Density

Use 1990 Census data to find the population density for your town or county or a few key census tracts, then compare that to the 2000 figures released in March.

Write a story about changes in the way your community is arranged. You'll find sprawl, sure, but where? Which direction is it moving? Is there less of a peak to the density but more places of moderate density? At the town level, you may find density falling simply because the kids have moved out of but their parents remain in the old homestead. You might find density rising where zoning has brought townhouses or apartments. You can do a micro-level version of this story at the census block level, which can be as revealing for a suburb of the 1920s or 1960s as it is for an 1880s rowhouse block.

Diversity

When 2000 Census data on race arrives in March, find the census tract or census block where the largest percentage of people declared more than one race. Ignore all community leaders, experts, interpreters and spinmeisters. Spend all day there for a week. Listen. Why are those people here? Who are they? Where is this place going? Then talk to experts. Write profile.

Contiuned on page seventeen

Continued from page sixteen:

Census stories

Undercount

Look at where the 1990 undercount was biggest in your area. (See *http://tier2.census.gov/pl94171/pl94index.htm.*) Use a spreadsheet to compare Census 2000 adjusted and unadjusted data to find the same thing. What's this place like? Do the people there think the undercount shows how they don't register in the eyes of the local government? Does the local government plan to adjust how and where it spends based on what's revealed by adjusted census data?

Overcount

There's some double-counting, especially in affluent suburbs where non-Hispanic whites predominate. (Check your richest census tract for a 1990 overcount at *http://tier2.census.gov/pl94171/pl94index.htm.*) When 2000 Census data comes out, use spreadsheet to find overcounted places, where adjusted totals are lower than the unadjusted totals. No politician likes to "lose" voters, so ask them what they think.

Migration

Get the latest annual IRS county-to-county migration data for your state. (See *http://ftp.fedworld.gov/pub/irs-soi/prodserv.pdf.*) Look at your county and see which counties it's receiving the most households from. Look at whether the comers have higher median income than the goers or vice versa. Use Census 2000 data to look at 1990s changes in racial makeup, etc., in those counties vs. your county. Use that insight to find new arrivals from those places and talk to them about why.

Immigration

Get the INS legal residency file (Available from NICAR's database library for as little as $55 for a year, see *http://www.nicar.org/data/ins/*). Find out who's coming from where, and which ZIPs they're congregating in. Use Census 2000 data for those ZIPs (translated to census blocks and tracts) to show how immigrants have (or haven't) changed that area's racial makeup, household and family structure, age, etc. And how big was the undercount in immigrant-heavy neighborhoods? See No. 5 above.

Boomers/centenarians/etc.

Single-year age detail by sex (next summer, state by state, watch/ask for Summary File 1, table PCT12) will let you look at any custom age group, like boomers, Gen Xers, draft-age men during Vietnam War, etc. It will tabulate how many of each year up to 100 years old, then people ages 101-104, 105-109 and 110+. Find the census tracts where they cluster (besides special-purpose group quarters tracts like nursing homes) and profile them.

Paul Overberg can be reached by e-mail at poverberg@usatoday.com

OFF THE LIST

Prison populations

A question posted to Census-L recently asked if federal prison populations are included in the 2000 Census data.

The answer is yes, that any "group quarters" population is counted within its respective geographical units. These include correctional facilities (local, state and federal), nursing homes, hospitals, psychiatric hospitals, schools for the physically handicapped, juvenile institutions, college dormitories, military barracks and ships, emergency shelters and group homes. These are just the highlights from a long list. The full list can be found in the layout for the SF1 data.

The SF1 data break down the populations of these various group quarters by race, ethnicity, sex and age.

In the PL94 data that was released in March, the only way to isolate a large prison that might be skewing a municipality's population is to identify the block number.

The SF1 data is expected to be released, state by state, between June and September. More suggestions about planning for this data will be included in *Uplink*'s next issue.

Uplink story ideas

Have you or one of your colleagues recently published a story using CAR that has not been done before or involved particularly difficult data work?

Do you know of a technical problem (or its solution) that others may like to hear about?

Or is there some issue or beat that we haven't covered?

If you have a story idea, we'd like to hear from you. Please contact managing editor MaryJo Sylwester by e-mail at maryjo@nicar.org or associate editor Amy Sherrill at amys@ire.org. Both can be reached by phone at (573) 884-7711.

Census and immigrants

By Paul Overberg
USA Today

Journalists interested in the local angle of foreign lands and people, can find less well-known data right now that gives surprisingly detailed information on immigrants in communities across the U.S.

Census 2000 data will provide information about immigrants too, but it won't be available for many months.

Current Population Survey

The single biggest and best source is the Current Population Survey, a 50,000-household survey that the Census Bureau conducts for the Bureau of Labor Statistics every month. While it's mostly done to calculate the unemployment rate, additional queries are made every one to two years about income, household and family structure, voting, education, school enrollment and childbearing.

And a whole new world of insight has been opened since the CPS began including in 1994 a question about citizenship on a variety of surveys.

For example, the most basic report on immigrants was published in January. It was the latest annual compendium on the U.S. foreign-born population, based on the March 2000 CPS. (See *www.census.gov/population/www/socdemo/foreign.html.*) It includes a simple eight-page overview, 60 data tables and links to comparable annual data tables back to 1995.

Citizenship question provides new data

And a whole new world of insight has been opened since the CPS began including in 1994 a question about citizenship on a variety of surveys. Consider this scenario: a perennial lament of political life is the declining voting rate, which is measured, among other ways, in the November CPS that follows the national election every two years. But a piece of this decline, however, is bogus because it's calculated by dividing the number of people who say they vote by the voting-age population. This may have worked fine in the 1960s, when citizenship rates skyrocketed after the U.S. opened immigration doors that had been closed for 40 years, but it isn't going to wash today.

New technique provides different data

Specifically, this country has millions of adults who can't vote because they're not yet citizens. So once the Census Bureau asked the citizenship question in its 1994 voting survey, it found that voting rates among citizens were several percentage points higher than measured among voting-age adults. In some states - notably California, Florida and Texas - rates are 5 to 10 percentage points higher than they were measured with the old yardstick. Predictably, national voting rates among Hispanics and Asians were 10 to 20 percentage points higher using the more accurate gauge.

The standard CPS reports don't always cross-tabulate citizenship by the subject focus, whether it's income, union membership or year of school enrollment. But it's there, and available through a tool called Ferret that lets you download or tabulate online the individual survey responses any way you please. (See *http://ferret.bls.census.gov/cgi-bin/ferret.*)

Limitations, however

One big limitation of the CPS is local focus. Standard reports cover national and regional totals, but there's not enough data strength in the survey for state or local levels. While you certainly can roll two or three years' survey data together to do that, (as the bureau does for its annual income reports), I would suggest talking with bureau experts before you try it with citizenship data.

That's one reason why Census 2000 data is so important. It not only provides more

Continued on page nineteen

Here are some other Census Bureau data that can help in international reporting:

•The bureau's international staff compiles population estimates and projections, by age, sex, fertility, work and marital status and other variables, for 227 countries and regions. The data is available in a variety of ways, including customized downloadable extracts and as animated population pyramids on the bureau's Web site at *www.census.gov/ipc/www/idbnew.html*).

•Interested in monthly foreign trade data? Global Trade Information Services, a Census Bureau vendor, takes the raw data and sells CDs and Web access with a variety of richer layers. Go to *www.census.gov/foreign-trade/www/index.html and www.gtis.com*

Continued from page eighteen: **Immigrants**

nuance on immigration and assimilation issues, but it's designed to show it for every neighborhood in the nation.

Details, details

So what questions are asked on the 2000 Census? Well, the queries parallel those on the 1990 Census, so there won't be comparability problems when it comes time to do stories. But these questions only are asked on the long form, which went to about one household in six.

That means you need to be a little cautious in interpreting the data, especially for very small areas. But the questions provide a wealth of detail that's worth the trouble. Each person is asked:

To really do focused analysis, wait for the Public Use Microdata Sample, which is a 1 percent or 5 percent sample of the actual census response forms.

• Ancestry or ethnic origin, such as Jamaican, Norwegian, Ukrainian, etc.

• Language spoken at home, and facility speaking English. This lets us analyze, by age and other variables, how isolated from English-speaking culture the people in an area are, and which languages they use. It also lets you find the neighborhoods with the most languages spoken at home.

• Country of birth. Among other things, this lets you see how clustered immigrants are, and how those clusters match – or don't – with concentrations of U.S. natives whose ancestry is from the same country. Do Irish immigrants cluster near old Irish neighborhoods? What kind of immigrants have filled that old Italian neighborhood on the south side of our city?

• Citizenship. Was the person born in the U.S., born in its territories, born abroad of U.S. parents, naturalized or not a citizen? How close together are the neighborhoods of highest and lowest citizenship rates? What's the national origin of the biggest groups in each one?

• Year of entry. Citizenship rates rise steadily with duration in the United States, but that will vary by nationality, education, etc. Why are my local XYZ immigrants naturalizing so much more slowly than the average for my area?

Check out

All this information can be analyzed for any area as small as a census tract (a neighborhood averaging 4,000 people) and broken down by all the other variables in the census. So you should consider:

• Checking how income and education affect things like citizenship rates by years since entry.

• Looking at whether, say, Salvadoran immigrants in your area naturalize at rates that vary from those of Dominicans and Guatemalans and Mexicans.

• Thinking about how you can use these questions to measure assimilation rates (speaking English, education, etc.) of various immigrant groups in a big city.

• Looking at neighborhoods with the highest rates of non-citizens and other languages spoken at home, and checking the difference in household and family structure from that of others in your city or metro area.

To really do focused analysis, wait for the Public Use Microdata Sample, which is a 1 percent or 5 percent sample of the actual census response forms.

It's not due out until 2003, and will be harder to work with and won't identify areas smaller than about 100,000 people, but it lets you do analysis you just can't do from standard census data. For instance, you could isolate just people born in the United States who don't speak English well, and analyze who they are by age, education, households they live in, etc.

Paul Overberg can be reached by e-mail at poverberg@usatoday.com

The following stories were written using U.S. Census data. To order these stories, call the IRE Resource Center, (573) 882-3364.

A 1991 *Des Moines Register* report investigated the high rate of infant mortality in Des Moines and found that half of the deaths were limited to just nine of the city's 52 census tracts. Story #8277

***The Miami Herald* dipped into the dating game by using 1990 U.S. Census data to find the percentages of singles in every zip code for two Florida counties. Story # 15786**

***The American-Statesmen* (Austin, Texas) conducted an analysis of car insurance records and combined that information with census data to discover that the people most likely to pay the highest insurance rates were those living in the predominantly minority and low-income zip codes. Story # 11065**

Computer-Assisted Reporting Boot Camps

Missouri School of Journalism
Columbia, Mo.

- **May 20-26, 2001**
- July 15-20, 2001
- August 5-10, 2001
- Jan. 6-11, 2002

More information, including a registration form, is available at *www.ire.org/training/bootcamps.html*, or call IRE at (573) 882-0684.

Joe Ellis' stories also will be available soon at the IRE Resource Center. Ellis of KMOL-TV, San Antonio wrote about his first venture in CAR reporting on page one.

Continued from page one:

Everyday methods

kept?" "How far back do those records date?" The whole process is always an adventure. And when we get the records we often find getting them was actually a simpler challenge than figuring out what to do with them.

Baby steps

Right out of bootcamp, my first CAR story didn't warrant much, luckily. They told us to start small and I think my colleague at The News of Texas, Paul Adrian, was going to make sure I did, anyway.

We took municipal court revenue reports from the State Office of Court Administration and found which cities were writing the most traffic tickets and making the most money per capita. Getting the data was as easy as a phone call and it was e-mailed to me a day later in an Excel spreadsheet. We did a little cleaning, a little crunching in Excel, and ran the story the next week.

We asked for this information about five different ways before we got what we wanted.

Unfortunately, soon after that The News of Texas announced its demise. For the next two months, my new CAR knowledge remained dormant, except for some minor tinkering I did at home. I can remember hearing the phrase, "use it or lose it."

Lying in wait for investigative opportunities to unfold, I helped an old friend in San Antonio find out who owed the city the most money for unpaid parking tickets. The friend, KMOL-TV reporter Brian Collister, had a spreadsheet listing the unpaid tickets that we analyzed to find Enterprise Rent-A-Car had a hefty tab. It wasn't long before I went to work with him at KMOL-TV, and I couldn't wait to dive into more data.

Collister already had some employee data from local school districts. Getting the same data for all school districts in the San Antonio area didn't take too long or cost too much. I put it all into text files -- nice and neat the way I liked it -- and imported those into an Access database. Then came the hard part.

I had a huge set of corrections data (that had to be wrestled away from the Texas Department of Public Safety, the state's law enforcement agency, for a not-so-nice price) that was so big and complicated I had to get some help from NICAR in cleaning and condensing it. This took a little time and a little money, but within a couple of weeks I had it back and ran it against the school employee data. Eventually, all of my time and work with this project wouldn't turn a news story, but it showed me what I could do by looking at records this way.

Missing Millions

After that, I verified some startling figures for missing school equipment in the San Antonio Independent School District. The initial tip on the story came from a "little bird" in the fixed assets accounting department, who told us that $1.7 million worth of equipment was missing over a three-year period.

We then requested the district to give us their list of missing inventory, however they declined to give it to us electronically and instead supplied the list on paper. The list totaled $2.5 million dollars of equipment lost over a five-year period. We asked for this information about five different ways before we got what we wanted. We had to narrow down what the school district had versus what we wanted, and what it was called because they weren't going to tell us what it was called and they didn't want us to report it.

I entered all the information into an Excel spreadsheet, added up the totals, calculated percentage increase/loss over time, sorted which schools were losing the most and looked at what items were lost the most frequently.

The list included things like computers, televisions, VCRs and band equipment. Two of the most unusual items were a John Deere tractor and a walk-in freezer.

Our story has caused quite a stir since it aired about six months ago. In early February, results of a district-wide audit of missing assets revealed $4.2 million worth of equipment missing.

Continued on page twenty-one

TRAINING

Stats in Chapel Hill

By MaryJo Sylwester
IRE and NICAR

One of the tools many journalists will use in reporting on the 2000 Census will be the USA TODAY Diversity Index, which measures racial and ethnic diversity and captures it in a single number.

How to use this index will be just one of the many techniques and concepts journalists will learn at the IRE and NICAR Advanced Bootcamp on Statistics in Chapel Hill, N.C. The six-day seminar will be held May 13-18 at the University of North Carolina.

Philip Meyer, one of the creator's of the diversity index and author of "The New Precision Journalism," will teach the seminar, along with *Washington Post* Database Editor Sarah Cohen.

This Bootcamp is designed for reporters who are comfortable with spreadsheets and database managers, and are ready to move on to statistical applications. The class is taught using SPSS, and includes an introductory lesson on the software.

Students will learn a variety of statistical tests, such as significance, association, correlation and basic regression; probability theories and sampling techniques.

Cohen shows examples of how statistical concepts were used in newspaper stories, and discusses how to explain it to readers in both the story and a nerd box. In the evenings, students can attend open labs to practice on datasets that journalists typically use, such as school test scores and home mortgage data.

Some of last year's bootcampers said the seminar does a good job of relating the concepts to journalism in practical ways.

"It's great to get trained on how to use SPSS by and alongside journalists, who will tend to use the same types of databases in the same ways on the same types of deadlines for the same sort of audience that you will," said Betsy Hammond, a reporter for *The Oregonian*, in Portland.

Dave Gulliver, who attended the bootcamp while working for the *Dayton Daily News* and now works for *The Virginian-Pilot*, said one of the benefits to the seminar is that Cohen and Meyer teach how to explain the concepts for both editors and readers.

"The bottom line is, two months later, you're more confident about what you can – and can't – do with stats," Gulliver said. "For us, the workshop paid off. Two months later, we ran a series on standardized tests, using linear regression to look at how a flaw in Ohio's school funding formula, which penalized poor schools, and affected their students' performance."

More information on this year's bootcamp, including a schedule, lodging information and registration form, is available at *www.ire.org/training/chapelhill2001.html*. Questions can be directed to NICAR at (573) 882-0684.

MaryJo Sylwester can be reached by e-mail at maryjo@nicar.org

Advanced Bootcamp Registration:
www.ire.org/training/chapelhill2001.html.

USA TODAY Updated Diversity Index:
http://cronkite.pp.asu.edu/census/
(Look under "What's New" for the link to "diversity index")

Statistics Books:
(from NICAR-L subscribers)
- "The New Precision Journalism," Philip Meyer (out of print)
- "SPSS Guide to Statistical Analysis," Maria Norusis
- "Seeing Through Statistics..." Jessica Utts
- "How to Lie with Statistics" Darrell Huff
- "News & Numbers," Victor Cohen (available from IRE Bookstore, *www.ire.org/store/books/*)

Note: A quick check at an online bookstore found all of these available, except Meyer's.

Continued from page twenty:

Everyday methods

Soon after, I discovered the biggest "slum lords" in San Antonio with the city's Code Compliance data, showing violations for the past 10 years. This database is maintained by the San Antonio Code Compliance Department, which inspects houses and buildings for code compliance and responds to citizen complaints about unsafe or unsanitary conditions.

The database was very dirty. Ultimately, I went through the 50,000 records, line by line, to make sure everything was typed in the same manner. Later, I obtained the agency's complaint database as well, and ran the two databases against each other. The complaints database had an "inspector comments" field that gave insight into the properties' conditions.

What I eventually found was that the most frequent violators owned many of their properties under various names. One turned out to be a doctor and part owner of the San Antonio Spurs NBA basketball team, who had properties under nine names (family trusts, business names, etc.).

Joe Ellis can be reached by e-mail at jellis@kmol.com

PHILLY

National Computer-Assisted Reporting Conference

OCTOBER 11-14, 2001 Philadelphia, PA

LEARN TO FIND FACTS ON THE INTERNET, DO DATABASE REPORTING ON DEADLINE, MAP CRIME, ANALYZE BUDGETS, PERFORM STATISTICAL ANALYSIS AND MORE.

CONFERENCE SITE: LOEWS PHILADELPHIA HOTEL

FOR HOTEL RESERVATIONS CALL: 215-627-1200

(HOTEL RESERVATION CUT-OFF IS FRI, SEPT. 14, 2001)

ROOM COST: $155 SINGLE/DOUBLE.

To register, please complete this form.
Visit our Web site at www.ire.org for the latest details.

REGISTRATION FORM

Please write carefully! This information will be used to make your nametag.

Name:_____

Employer/affiliation/School:_____

Address:_____

City, State:_____

Zip Code:_____-_____ E-mail:_____

Phone:_____ Fax:_____

TO ATTEND THIS CONFERENCE, YOU MUST BE A CURRENT IRE MEMBER THROUGH 11/1/01. MEMBERSHIPS ARE NON-REFUNDABLE.

_____($150) I'm an IRE professional member and would like to attend the conference Oct.11-14.

_____($100) I'm an IRE student member and would like to attend the conference Oct.11-14.

_____($200) I would like to attend the conference Oct.11-14 and need to join or renew my U.S. membership.

_____($205) I would like to attend the conference Oct.11-14 and need to join or renew myinternational membership.

_____($125) I would like to attend the conference Oct. 11-14 and need to join or renew my student membership.

_____($25) Late fee for registrations postmarked or faxed after Sept.17.

_____Total

To register, mail this form and a check to NICAR, 138 Neff Annex, Missouri School of Journalism, Columbia, MO, 65211. To register by credit card, you must have a Visa or MasterCard. We cannot accept American Express. You may fax your credit card registration to 573-882-5431 or register online at www.ire.org/training/philly.

Card Number:_____ Expiration Date:_____

Card Holder Name:_____ Card Holder Signature:_____

Cancellations must be sent via e-mail to jgreen@ire.org. There is a $50 processing fee for all cancellations until Oct. 10, 2001. Refunds will not be given for cancellations after Oct. 10, 2001.

Registration form

REGISTRATION

To register, please complete this form.

Visit our Web site at **www.ire.org** for the latest details.

Please write carefully! This information will be used to make your nametag.

Name: _____

Employer/Affiliation/School: _____

Address: _____

_____

City, State: _____

Zip Code: _____–_____ E-mail: _____

Phone: _____ Fax: _____

To register, mail this form and a check to IRE, 138 Neff Annex, Missouri School of Journalism, Columbia, MO, 65211. To register by credit card, you must have a Visa or Mastercard. We cannot accept American Express. You may fax your credit card registration to (573) 882-5431 or register on-line at www.ire.org./chicago/.

Cancellations need to be via e-mail to jgreen@ire.org. There is a $50 processing fee for all cancellations until June 13, 2001. Refunds will not be given for cancellations after June 13th.

Card Number: _____ Expiration Date: _____

Card Holder Name: _____

Card Holder Signature: _____

To attend this conference, you must be a current IRE member through 7/1/01. Memberships are non-refundable.

MAIN CONFERENCE DAYS

_____**($150)** I'm an IRE professional member and would like to attend the conference June 15-17.

_____**($100)** I'm an IRE student member and would like to attend the conference June 15-17.

_____**($200)** I would like to attend the conference June 15-17 and need to join or renew my U.S. membership.

_____**($205)** I would like to attend the conference June 15-17 and need to join or renew my international membership.

_____**($125)** I would like to attend the conference June 15-17 and need to join or renew my student membership.

_____**($25)** Late fee for registrations postmarked or faxed after May 21st.

CAR DAY (optional)

Thursday, June 14 is the optional Computer-Assisted Reporting Day and requires an additional fee. If you would like to attend, please check one of these options:

_____**$50** I would like to attend the CAR Day June 14 as a professional or international IRE member.

_____**$35** I would like to attend the CAR Day June 14 as a student IRE member.

_____Total

Bits, Bytes and Barks

NICAR Data Updates

The NICAR Data Library recently updated the Federal Aviation Administration Accident and Incident database and the MAUDE database on medical devices.

The MAUDE database, which stands for Manufacturer and User Facility Device Experience, includes reports of problems with medical devices from 1991 through the end of 2000. Health officials and product distributors and manufacturers report adverse effects with medical devices to the U.S. Food and Drug Administration.

The FAA Accidents and Incidents database includes all reports from 1990 through about mid-2000. Not all reports from 2000 had been entered into the database by the FAA when NICAR requested the data. Earlier data going back to 1974 is available separately from NICAR.

Cost: $40, $80, $105, depending on the size of your news organization.

For more details about either dataset, including sample slices and record layouts, go to *www.ire.org/datalibrary/datasets*. To order, call (573) 884-7711.

IRE National Conference

The 2001 IRE National Conference will be held in Chicago, June 14-17, at the Hyatt Regency Chicago.

The first day of the conference will be an optional CAR day, featuring discussions on the latest in computer-assisted reporting.

Panels on the optional CAR day will include: Insurance: The data and stories we're missing; Health Care: From doctor data to hospital data; Teaching CAR: For trainers and educators; Building Your Own Databases; School Data: Tests, budgets and scams; CAR for Broadcast: Best uses; Census – Using the Census for Non-Census Stories; Digging into Votes: Nuts and bolts lessons from Florida.

Cost: $50, $100 or $125, depending on the size of your news organization.

In addition to the wide array of panels on investigative reporting topics, the conference will also include a job fair, reprints of some of the best investigative work of the past year, a TV show and tell room and a computer demo room.

Registration forms and more information is available on the IRE Web site at *www.ire.org/chicago/index.html*

Campaign Finance Data

One way to stay ahead of upcoming elections is to have constant access to campaign finance data. NICAR provides federal data from the Federal Election Commission on a monthly subscription basis. That means you'll receive a CD-ROM with the latest data at the beginning of every month.

The data is only for federal candidates. It includes contributions from invididuals and committees, as well as transactions from one committee to another.

Cost for the annual subscription is $100 for circulation under 50,000 or 50-200 market; $200 for circulation 50,000 to 100,000 or 25-50 market; and $225 for circulation over 100,000 or top 25 market.

More information, including sample data and record layouts, is available at *www.ire.org/datalibrary/databases/feccc/* or by calling the NICAR Database Library at (573) 884-7711.

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

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