Uplink

April 1998

A newsletter for the National Institute for Computer-Assisted Reporting

VOTER FRAUD

Errant electioneering

By Dan Keating
The Miami Herald

When the election results for mayor of Miami came in on the night of Nov. 4, it seemed obvious that something strange had gone on in the voting, the first for the new "strong mayor" form of government that would give the city's chief executive a new level of power.

Incumbent Joe Carollo won a majority of the votes cast at the polls. But he was trounced by a 2-1 margin in absentee ballots by one of his challengers, former mayor Xavier Suarez. The absentee tide stopped Carollo 155 votes

---

TECH TIP

Enter Intranets

By George Landau
NewsEngin Inc.

Who among us wouldn't want a newsroom where reporters can search public records in a flash, without a lot of training and without having to leave the comfort of their ergonomic chairs?

All it takes to fulfill the promise of an Intranet are a newsroom full of modern PCs or Macs, a network with Transport Control Protocol/Internet Protocol (TCP/IP), a Web server and a big chunk of your time.

There's no way around it: setting up an Intranet for CAR is a lot of work. But if you make the right choices at the outset, your work can pay large dividends. As someone who has been wrestling with this stuff for a few years, I can offer some advice on getting started.

The network

Since an Intranet uses the same technology as the Internet, your network

Continued on page four

---

short of a outright election, forcing a runoff that he then lost.

The strange voting pattern and unprecedented number of absentees – more than 10 percent of the ballots cast – prompted Carollo, the Florida Department of Law Enforcement and *The Miami Herald* to launch investigations into what had gone on.

The Herald has tracked down and published more than 300 "problem" votes so far:

• A dead guy.

• People who don't live in the city or live in different city commission districts than where they voted.

• People who swore they didn't vote, despite a record showing a ballot cast in their names. Many said their names had been forged.

• People who said ballots were taken from them.

• People who said they were paid $10 to vote absentee.

• People who were solicited to vote by a former food stamp eligibility officer who still worked as a facilitator for elderly immigrants.

• People whose absentee ballots were "witnessed" by people who never saw the voter or by false signatures.

• More than 100 felons who had lost their civil right to vote.

Contacting more than 1,400 voters required a unique shoe-leather effort that lasted more than two months. But the project also involved a heavy dose of computer-assisted reporting to set the plate for the reporters fanning out into the community, organizing what they found and painting the emerging portrait of fraud.

The essence of the computer work – to find potentially suspicious activity – was matching different lists against the list of people who voted in the election. Though

Continued on page two

THE ROAD FROM INDY

Update

Local elections are often more colorful than their national counterparts. They are just as often susceptible to abuse. Dan Keating of *The Miami Herald* reviews his paper's extensive probe into voter fraud committed in the city's mayoral race. Local law enforcement has followed in the footsteps of the *Herald*'s investigation.

This issue highlights the drive toward installing Intranets in newsrooms. NewsEngin's George Landau offers a Tech Tip on setting up these systems. NICAR database administrator Justin Mayo surveys the CAR specialists on the front lines of applying the technology.

Intranets weren't the only hot topic at Indiana CAR. NICAR's Seth Hemmelgarn takes the pulse of an apparent software shift to Access but gives recognition to the FoxPro faithful. And CFIC's Jack Dolan gathers tips for discerning which campaign contributions merit making the news.

PAGE 10
Stats from the Road
PAGE 11
Handout of the Month
PAGE 12
On the Internet
PAGE 14
First Ventures

Continued from page one: # Problem polls

Florida law specifically forbids releasing the voting list to the public, *The Herald* obtained it.

With that voter roll, we could wrest great value out of the other public records that we keep on hand.

Movers and matchmakers

To start off, we matched the voter list against the Social Security Master Death List we had obtained from NICAR, though that list is somewhat spotty by having fields for ZIP codes but not full addresses. Since the voter list did not include a social security number, we had to match by name, date of birth and ZIP code. To broaden our search, we matched name and date of birth to anyone whose last ZIP code was in our county.

It looked exciting – we had eight exact matches. But, much to our disappointment, our field reporters found the purportedly dead voters quite alive and quite well: each proved that the rumors of their deaths had been greatly exaggerated. Most told tales of fighting with the Social Security bureaucracy about being listed as deceased.

(We learned from reporters around the country that such errors aren't at all unusual. And the one dead man we did identify didn't appear in the Master Death List.)

Packed voting houses

Our next step was more straightforward. With the list of 44,000 voters, we simply grouped by address and sorted descending to find the addresses housing the most voters. Since addresses included apartment or unit numbers, the search aggregated individual living units rather than apartment buildings and showed many residences with six, seven, eight or more voters.

To make the list of voters per home more useful, we merged the addresses with the county's property appraiser roll, which we obtain each year for real estate analysis. Matching the addresses didn't go smoothly because the two agencies are inconsistent when it comes to "3 Ave." or "3rd Avenue," and "S.W. 29th Ter." or "SW 29 Terr."

Using SAS for Windows, we cleaned the addresses from both lists. We removed all punctuation and letters ("th," "st," and "rd") from the fields with numeric street names and scanned to replace variations of "ave," "avenue," "av," and so on.

Matching with the property roll allowed us to identify single-family homes with one or two bathrooms and more than five voters. It also showed us vacant and non-residential properties with voters. When we visited the properties with many voters, we found out-of-town and phantom voters stashed there.

To confirm that people lived elsewhere, we used driver's license and vehicle registration information through AutoTrak. We also used homestead exemptions, a tax deduction that can be taken only on your primary residence. We found people whose driver's license, vehicle registration and homestead records indicated that they lived outside the city. In every case, we spoke directly to the person or, if that was impossible, sent a letter asking for proof of being a legitimate voter.

The cell bloc

One seemingly straightforward match proved tricky: matching voters with the list of people sentenced to state prison or probation in Florida. The list, available on CD-ROM from the state Department of Corrections, doesn't include those sentenced to time or probation locally, but it captures a good population of convicted felons, who have lost their right to vote.

Because the prison list doesn't have an address field and the voter roll doesn't have a social security number field, we matched by first name, last name, middle initial and date of birth for people convicted in our county or neighboring counties. Of the more than 400 matches, almost none were false matches caused by two people having identical names and birthdays. So it looked like we had a lot of illegal voters.

Complications arose when we checked court files for our first batch of interesting felons. Almost none, it turned out, had been convicted. Miami's overloaded court system is notorious for convicting rarely. Almost everyone can reach a plea agreement with

Ballots for hire

"adjudication withheld." About three-quarters of the people sentenced to state probation had not been convicted. People who would probably be convicted elsewhere in the state and lose voting rights had not lost them in our county.

We accessed online the local criminal court docket to find which of the 400 offenders had been actually convicted. We determined that more than 130 had been, but the search still wasn't over. We ran those 130 names past the state's Executive Clemency Board that handles restoration of rights. Sorting through the index cards where they keep their records, we found more than 100 felon voters who did not have their rights restored.

We also compared the list of people who voted with the entire list of registered voters and looked for people who had changed addresses into the city or between city commission districts near the election. Under Florida's implementation of the motor-voter law, you can announce on election day that you've moved into the district. And you're allowed to vote.

Available absentees

Computer matching helped us draw webs of conspiracy involving the organizers of bad votes. A database of people who witnessed absentee ballots showed that many campaign workers witnessed dozens of ballots. If any witnessed ballots that we knew were bad, we checked other ballots handled by them.

We checked bad ballots against a database of campaign contributions and expenditures we had made from paper records for stories before the election. We found people who used their real, out-of-city address to contribute but an in-city address to vote.

Mapping uncovered the operation that paid people $10 to vote absentee in the runoff election. Using ArcView 3.01, we found participants and proved the scope of the operation being run from a parking lot in a poor neighborhood.

In that neighborhood, the primary election included the mayor's race and a contested city commission race involving a popular candidate.

People had months to obtain and fill out an absentee ballot for the election. In the runoff, voters had less than a week to obtain and return an absentee ballot. By mapping all absentees within three-quarters of a mile of the parking lot in both races, we discovered unlikely results: Dozens more absentees were in the runoff than in the main election. By comparing the list of voters from the two elections, we found people who had troubled themselves to vote absentee in the runoff but had not bothered to vote at all in the primary election. We targeted that list s in our search for paid voters. Several confessed, and others admitted knowing of the operation but denied that they were personally involved.

Maiden Intranet voyage

To keep track of our contacts of more than 1,400 voters, reporters carried paper forms to collect consistent information. Results were then entered in an Access database. As the project progressed, the Access front-end was replaced by a Web-based version that christened our Intranet.

We used the Intranet to set up Web-based search forms so everyone could look up reporter notes on voter contacts. Reporters and editors could search lists of people who had voted or witnessed absentee ballots. Empowering team members to handle their own look-up chores was instantly popular and prevented a bottleneck at the geek's office.

After a three-week, non-jury trial, a judge in March overturned the election and called for a new vote. An appellate court then went a step further by putting Carollo back into office and declaring that the voting-fraud participants should be punished rather than simply forced to compete in another election.

The Florida Department of Law Enforcement has followed our tracks frequently, arresting the man we identified as running the paid-voter scam and several of the illegal voters we uncovered. More arrests are pending.

The archive of stories from this project is available at www.herald.com/dade/archive/elect97/index.htm

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

Newsroom boost

needs to use the Internet's TCP/IP networking protocol. When your Intranet users fire up their Web browsers, they'll use TCP/IP to connect to your internal Web server. Chances are pretty good that your network already has TCP/IP, but you should make sure before diving in.

The Web browsers

If possible, make sure that all of the potential users of your Intranet are using the same kind of Web browser and that it's Version 3 or later. Netscape 3 and 4 and Internet Explorer 3 and 4 are all fine choices. Both companies' browsers are free these days, but my favorite on both Windows and the Mac OS is IE 4.

It's not essential that everyone use the same browser software, but it'll make your job easier if they do. That's because when you design a Web page, you'll discover that it can look different depending on the browser with which you view it. While it's best to design pages that look great in both Navigator and IE, this can mean more work for you.

The Web server

Assuming you already have a TCP/IP network and that everyone on it already has a decent Web browser, the real work will be in setting up your Web server and putting some useful data on it.

If you're not just dabbling and want a robust Web server, I strongly recommend you use Windows NT Server for the operating system and Microsoft's Web server software, Internet Information Server. Microsoft released IIS 4.0 a few months ago, and it's an excellent product. It's fast, powerful, widely supported and free. You can download it as part of the Windows NT Option Pack from *http://www.microsoft.com/ntserver*.

If you can't afford NT Server (about $500), the next best thing is NT Workstation (about $200). NT Workstation has the same underlying code as NT Server, so it's just as fast and just as dependable. You can download a version of IIS for NT Workstation called Personal Web Server, also part of the NT Server Option Pack.

If your newsroom can't afford NT Work-

station and you can't afford to get a job in another newsroom, you can install Personal Web Server for Windows 95. This will work fine in the beginning, but Windows 95 is simply more prone to crashes and does not run as fast as NT. If you want to get your colleagues hooked on your Intranet, it's important that you build a system that runs well with a minimum of downtime.

Regardless of which Microsoft operating

You should plan to put all of your Web-searchable data on the Web server's local hard drives. Do not keep the data on a network drive, as this will lead to pokey searches that bog down the entire network.

---

system you use, make sure the server has plenty of RAM, especially if you intend to design Web applications that search databases. I'd start with at least 32 megs of RAM. If you plan to use databases, go for 64 megs or more. RAM is cheap these days.

In terms of the speed of your CPU, I'd recommend at least a Pentium 133 for your Web server. Faster is definitely better. Fast hard drives are also a good way to boost the performance of your server.

In terms of the hard disk space you'll need, this depends on how much data you plan to share. Set aside about 500 megs for the operating system and applications (planning ahead for the inevitable code bloat in Microsoft's future offerings).

Unless you'll be using a client/server database like Microsoft SQL Server or Oracle, you should plan to put all of your Web-searchable data on the Web server's

Accessing data

local hard drives. Do not keep the data on a network drive, as this will lead to pokey searches that bog down the entire network.

Building the Web pages

When it comes to building Web applications that search a database, there are basically two popular ways to go. Which one you choose depends on your budget, ambitions and prior experience. Either approach will require that you learn the basics of HTML forms and tables.

• Approach #1:

Allaire's Cold Fusion *(http://www. allaire.com)*

Cold Fusion (about $500) extends HTML with special tags that support straightforward SQL queries. It includes a well-designed wizard that does most of the work in building basic search forms and result tables. You can tweak the wizard's offering using any text or HTML editor.

Cold Fusion works in partnership with Microsoft's IIS (and other Web servers) by acting as a broker for database requests. When a user submits a query, the Web server passes the request to Cold Fusion, which runs the query using your database software and passes the HTML-formatted results back to the Web server, which in turn sends it back to the user. The whole process is pretty fast.

• Approach #2:

Microsoft Active Server Pages

ASP, which is built into IIS and thus costs nothing, is essentially a version of Visual Basic (VB) that runs on the Web server. To use it, you write Web pages that contain a combination of Visual Basic programming and HTML. Your users and their browsers never see the VB part – they get just the HTML. Just as with regular VB, you can write ASP pages that query databases and manipulate the results.

If you intend to build a moderately complex Web application with conditional logic and loops, ASP might be a better choice than Cold Fusion. If you already are comfortable with VB or another flavor of BASIC, you'll be able to scan the ASP documentation and get up to speed pretty quickly.

ASP works in a similar fashion to Cold Fusion, but since ASP is more tightly integrated with the IIS Web server, it runs a little faster. For normal newsroom use, though, this speed difference isn't noticeable. The speed of a Web database application depends largely on your choice of hardware and database software, not on your choice of ASP or Cold Fusion.

If you go the ASP route, it's probably worth spending about $200 for Microsoft Visual InterDev, which includes wizards that build some basic database applications for ASP. InterDev does a nice job of managing the various components of a Web application, and it includes a version of the FrontPage HTML editor, so it's worth the money. A new version of InterDev is due soon and is supposed to provide much-needed improvements in the visual design tools (in the current version, "visual" pretty much means you're allowed to see your Visual Basic and HTML code as you write it).

Managing the data

Active Server Pages and Cold Fusion both use ODBC (Open Database Connectivity) to access your data. If your database software provides an ODBC driver, these tools should work for you. Access, FoxPro, SQL Server and Oracle all have ODBC drivers that work well for basic Web queries. It's a good idea to make sure you have the latest ODBC driver for your software by downloading it from the database vendor's Web site and installing it on your Web server.

As for which database software is best suited for newsroom Intranets, that's a question for another Tech Tip.

But if you want folks to be able to use more than a Web browser (Access or Excel, for example) to work with large databases on the server, you should consider setting up a client/server system like SQL Server.

If you'd rather focus on Web users and don't want to mess with client/server, FoxPro is probably faster than Access and has fewer limitations on database size. Tom Boyer of *The Seattle Times* has succeeded with an Intranet based on FoxPro, Cold Fusion and Windows NT.

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

Conference recap

Indiana CAR showed not only how far the skills of journalists have improved during the last year, but what new skills they are seeking. This year more journalists than ever signed up for classes in mapping(ArcView) and statistics (SPSS). As you will see below, the demand for tools to share data in the newsroom was at an all-time high.

And if you missed Indiana CAR, watch for upcoming seminars that will offer more concentrated training in these topics.

Learning to share

By Justin Mayo
NICAR

Judging from the activity at Indiana CAR, the next challenge facing computer-assisted reporters will be navigating the quagmire of Intranets. Sounding more like a spelling bee than a journalism conference, terms such as TCP/IP, ODBC and ASP were heard throughout the halls of the Hyatt. And although it's easy to get lost in the technical aspects of setting up an Intranet, there is a simple, journalistic motivation driving CAR specialists and database editors to immerse themselves in acronym hell – sharing data with the rest of the newsroom.

"A reporter has easier access to a government computer thousands of miles away than a colleague's computer three feet away," said Tom Torok of *The Philadelphia Inquirer.* "In such an environment, each computer and journalist is an island. However, the information could be made available to all, quickly and securely, through client/server arrangements."

At the conference, Torok, George Landau of NewsEngin Inc., Tom Boyer of *The Seattle Times,* and Bill Miller of the *Battle Creek Enquirer* led roundtable discussions on the choices available in setting up an Intranet.

"Journalism is about answering questions quickly and authoritatively, and a well-stocked client/server system helps reporters do just that," Landau said. He pointed out that these systems provide the best way to keep large amounts of clean, reliable data – such as voter registration, property, and corrections records – readily available for deadline queries.

Indiana CAR marked the first time a NICAR conference has offered such exten-sive coverage of Intranet technology. The response was enthusiastic.

"I was shocked at the number of people we had in our session," said Boyer. "This field is changing so rapidly.

As the demand for client/server systems in newsrooms grows, so too will the necessity for CAR specialists to learn the technology.

"The classes and panels gave folks a taste, but I think they wanted more," Torok said. "In the future, I'd like to see an all-day workshop where people can bring in samples of their databases and walk out with a finished application."

But will the trend toward Intranets push CAR specialists into the role of full-time propeller heads? Most people argue that

---

"A reporter has easier access to a government computer thousands of miles away than a colleague's computer three feet away."
– Tom Torok, *The Philadelphia Inquirer*

---

although the impetus for setting up an Intranet will have to come from the newsroom, the long-term goal should be to get the information-systems experts involved.

"I think the ideal situation might be that you have technology-savvy journalists who can oversee these projects, but full-time database wonks, whose skills are way beyond journalists', to set up and maintain the systems," Boyer said.

Landau emphasized that cooperation between the two departments is possible "as long as (information-systems) people are working for, not against, the newsroom." Client/server realms such as data cleanup and verification, he said, should remain newsroom chores because they involve journalistic judgment.

Continued from page six:

In the long term, assuming the technical hurdles of setting up an Intranet don't consume all your energy, you should have more time for the real work of journalism. If all reporters can quickly and easily access the information they need, then CAR specialists may spend more time on hard-core analyses.

"Even if the resident CAR guru has to do most of the work, setting up the server can still save time in the long run because it allows reporters to do more digging on their own," Landau said.

Once a client/server system is in place, Boyer and Torok stress that all kinds of information can be shared. *The Seattle Times*, for instance, is experimenting with free-text queries over its Intranet. Not only can structured databases be shared throughout the newsroom, but Boolean searches of huge text blocks can be run.

"Don't just think of 'data' as in databases," Torok said. "Just about any form of information - spreadsheets, text, graphics, even Power Point presentations - can be manipulated and made searchable to suit your staff's needs."

FoxPro to Access and back again

By Seth Hemmelgarn
NICAR

Looking at the class schedule at Indiana CAR, one trend stood out:

Most classes were taught using Access, not FoxPro. This reflected the shift in newsrooms across the country; use of Access has become more widespread.

"Ever since Access 2.0 came out and proved to be fairly stable, there's been a steady drift toward deploying it for general newsroom use," said NewsEngin's George Landau.

As part of the Microsoft Office Package, Access is readily available, and newsrooms just starting CAR programs – even those already having them – are using Access more frequently.

If you don't know anything about Structured Query Language, Access' query grid can be a simpler way to build queries. You can point and click your way through them. But if you do know SQL, the grid can make analysis harder.

"I'm still resisting," said *The Cleveland Plain-Dealer's* Beth Marchak, who prefers FoxPro's more direct approach to analysis. "I think in SQL. Learning Access is like being forced to walk around the barn to gain entry when I'm already standing at the front entrance."

"Learning basics in Fox makes it easier for anyone to apply basic concepts and database skills no matter what program is used."
– Dave Fallis,
Tulsa World

Marchak also shies away from Access because she doesn't want it choking on her data, as previous versions of the software have been notorious for doing. "What good is a database manager that can't handle more than 250,000 records?" Marchak said. "All my databases are much bigger than that."

Though maybe not the most fashionable software, FoxPro is still worth learning. "Learning basics in Fox makes it easier for anyone to apply basic concepts and database skills no matter what program is used," said the *Tulsa World's* Dave Fallis.

FoxPro's library of string functions, downloadable utilities, and easily understood programming/scripting language should keep the program from getting too antiquated, Tom Boyer of *The Seattle Times* said.

Landau uses Access to keep track of queries and build front-ends, but he still prefers FoxPro for processing and cleaning. "Even Access97 continues to lack many of the important data-cleaning functions that are built into FoxPro," he said. "It does seem, however, that Access97 is considerably faster than its predecessors at running large queries."

Access does have its advantages. It has more export options. You can save an Access file as a FoxPro file, but not vice versa. Fallis also finds the import wizard handy for quick

and dirty imports into blank database structures. Working with relational databases is where Access has the most obvious advantage over other software. "FoxPro is not and never will be a purebred relational database," Landau said. "Access, in contrast, has relational integrity and validation rules hard-wired into its database engine."

Landau assumes this partially explains why Access is slower than FoxPro at some tasks: it takes a no-frills approach to data manipulation.

With the proliferation of Access, what will become of FoxPro and others?

Landau, for one, isn't sure. "Once SQL Server is made less scary and more affordable, I bet they'll try to wean everyone off of FoxPro and onto a single platform." Landau's not thrilled with this idea, since SQL Server is slower than FoxPro at reformatting large amounts of data. But, he said, SQL Server beats both Access and FoxPro at allowing lots of people to query the same database simultaneously.

Despite Access' popularity, those who prefer FoxPro may never give it up completely. As Boyer said, "I'll probably still have FoxPro on my machine when I'm old and gray."

Newsworthy contributions

By Jack Dolan
CFIC

If the CEO of a casino company gives $1,500 dollars to a state senator running for reelection, is it news? Giving money to a candidate you believe in is not undemocratic. For most busy people, it is more practical to promote their ideals by writing a check than by stumping door-to-door with brochures. So when do campaign contributions become news?

The consensus at Indiana CAR was that determining which contributions lead to good stories is difficult to do with contribution data alone. Experienced campaign data crunchers, like David Poole of the Virginia Public Access Project, suggest taking the analysis further. First, Poole says, get the candidate's financial disclosure forms. If the candidate has two kids working for a casino or owns stock in a casino, noting a few hundred dollars' worth of gaming company contributions barely scratches the surface of the story.

Next, check the committee assignments of incumbents. Kevin Murphy of *The Kansas City Star* recently wrote about contributions of $1,500 from gaming executives to the co-chair of the gaming and wagering committee of the Missouri Senate. The senator recently proposed abolishing hourly loss limits for casino customers.

Then there are the aggregate stories: Which industries gave the most? What did they get in return? These stories are the real data challenge because summing contributions from a single contributor can be difficult if the name is entered in different ways. The difficulty grows exponentially when you try to sum contributions from an entire industry.

But if you have the patience and data skills to standardize contributor names, the reward can be staggering. With technical help from the Center for Public Integrity, *The Indianapolis Star/News* chronicled how a consortium of businesses aligned with the Indianapolis Chamber of Commerce went on an unprecedented spending spree and, coincidentally, got its entire legislative agenda passed. According to the *Star/News* report, "When some lawmakers tried to protest or even debate the issues, they were reminded in private meetings to stick with the party – and the money that got them elected."

Another good aggregate story, by Shirish Date of *The Palm Beach Post*, clawed its way to the heart of one of the sadder ironies of the modern legislative process. *The Post* requested records of contracts with registered lobbyists from every public agency in the state and then entered those records into an Excel spreadsheet. According to the resulting story, "Last year, Florida taxpayers paid more to lobbyists to influence their own legislators than they paid in salaries to those same lawmakers."

Wendell Cochran of American University reminds reporters to understand the system the data describes. "For every complex problem, there is a simple answer that's wrong." Cochran says one oversimplified approach to campaign finance reporting is the bribe paradigm: rich businessmen bought candidate X's vote. Cochran says we should not ignore the possibility that candidate X solicited the rich businessman's contribution. He calls this the "blackmail paradigm." Cochran advises asking contributors how they came to give the money and getting on the candidate or committee mailing lists so you know when they are directly soliciting from industry-wide groups.

IRE NATIONAL CONFERENCE
JUNE 4-7, 1998
NEW ORLEANS

Come to the Big Easy for a first-rate conference in a first-class city. For more information on the 1998 IRE National Conference call us at (573)882-2042 or visit IRE's web site www.ire.org.

Conference Site:
The Fairmont Hotel
For Hotel Reservations call:
1-800-527-4727
Room Cost:
$119 (single) or $139 (double)

For more information,
call 573-882-2042
or see the IRE web page at
www.ire.org

REGISTRATION FORM

Please write carefully! This information will be used to make your name-tag.

Name: _____

Affiliation: _____

Address: _____

_____

City, State: _____

Zip Code: _____ - _____ E-Mail: _____

Phone: _____ Fax: _____

To attend this conference, you must be a current IRE member through 7/1/98. ____
Please indicate your membership status:____

_____ I am a current IRE member through 7/1/98
_____ ($40) I need a new professional membership or to renew my professional membership
_____ ($25) I need a new student membership or to renew my student membership
_____ ($55) I need a new intn'l membership or to renew my intn'l membership
_____ ($150) Conference registration fee
_____ ($50) Optional CAR day registration fee
_____ ($50) Conference registration fee for one day only; please indicate day:_____
_____ ($100) Student conference registration fee
_____ ($35) Student optional CAR day registration fee
_____ ($15) Late fee for registrations postmarked after May 22, 1998
_____ Total

To register, mail this form and a check to IRE, 138 Neff Annex, Missouri School of Journalism, Columbia, Missouri, 65211.

To register by credit card, you must have a Visa or Mastercard. We cannot accept American Express. You may fax your registration to (573)882-5431 or register on-line at www.ire.org.

Visa or Mastercard #:_____ Exp. Date:_____

Card Holder:_____

Card Holder Signature:_____

PRELIMINARY SCHEDULE
IRE NATIONAL CONFERENCE
IN NEW ORLEANS, JUNE 4-7

IRE will be offering 15 panels on computer-assisted reporting on Thursday, June 4, with experts from around the country.

On Friday and Saturday, there will be 50 panels on better techniques in reporting and editing, covering beats, specialites, projects and international reporting. The speakers will include leading investigative journalists and recent prize-winners, in both print and broadcast.

Both Friday and Saturday will have fast track, 50 minute panels on special skills. There will also be a special broadcast track both days. In addition three tracks each day of main panels, each lasting nearly an hour and a half, will be offered both days.

Two showcase panels will be presented, one on early Friday evening and one on Saturday evening.

One will feature long-time investigative reporters Seymour Hersh and Jack Newfield. The other will feature book authors and

reporters examining the coverage of ethnic issues.

On Friday night, *The Times-Picayune* will host a reception at the city's spectacular aquarium. At the Saturday luncheon, IRE will present its contest awards to the winners on Saturday at 4:30 p.m., IRE will hold its annual membership meeting and elections.

The Poynter Institute will lead three simultaneous, two-hour sessions on Sunday morning. A complete program with the list of panels and speakers will be out soon.

Causes and effects

By Sarah Cohen
NICAR

One stumbling block that reporters face when they begin working with statistics is the timeless cause-and-effect debate.

When virtually every book on numeracy and statistics lectures us about our constant confusion between correlation and causation, it may seem wise to avoid any analysis that will thrust you into the debate.

But complaints are often a little hysterical. I doubt many of us would write stories ascribing (a) more drinking to more clergymen in town, or (b) higher stock prices to sun spot activity.

Even so, the point is well taken. It's all very nice to say that the failure to get a mortgage or the amount of bail set are associated with a person's race. Those only become stories when we can say why. We aren't interested in just linking two events; we want to show why they're related.

Statistical tests such as chi-squares and regression results can tell you whether a relationship between two events could have happened by dumb luck. But they can't tell us whether one event caused another.

More bad news: We rarely conduct controlled experiments, the only kind of research that can tell you cause and effect. Instead, we work from a mass of public records that were collected for some other purpose. These records often miss the key elements and defy our attempts to control for even simple and obvious confounding events.

School reports, for instance, lack the single biggest indicator of poor performance – a definitive measure of poverty. Home mortgage records lack any information on credit records. And health records lack medical histories.

Validating research

At least one author is ready to help us through the steps that others take when they want to say that one event may cause another. In Statistics: A Spectator Sport, author Richard M. Jaeger reveals the ways researchers decide whether they've found a cause and effect.

We usually only care about what researchers call "internal validity" of our work. External validity would mean that we care about generalizing from our specific work to a bigger world – something we rarely do. So here are some of the guidelines:

• Does the cause come before the effect?

This seems obvious, but it isn't always. Test scores may depend on the number of teachers in a classroom. The timing, though, is relative. Spending may, in turn, depend on performance just last year. Performance in a school doesn't usually change that quickly.

• Is there a third event that causes the relationship? Again, a school example is worth noting. Most reporters would find that high minority populations in schools are associated with low test scores. But a third factor – poverty and education of the parents – plays a role in both the cause and effect in this case. Although ethnicity comes before just about anything else, few (outside *The Bell Curve* crowd) would argue that it causes any of the difficulties minorities face in school.

• Is there another explanation? Can you think – or have your sources thought up – of rival theories that are equally sound? Look to your reporting effort rather than statistics for this step.

• Does anything refute the theory? Are there other studies that fail to show the expected causal relationship? Why might your theory only work locally? Why might it fall apart in another year?

This is an uncomfortable position for us. It involves an about face in our normal reporting process – we're faced with disproving our own theories rather than proving a tip. And it leaves us open to criticism we've never faced when collecting quotes or uncovering documents.

But there is good news. One way to figure out cause is to watch and listen for anecdotes that are described by the numbers.

Now that's something we're happy to do. In fact, we prefer it. We'll search for weeks to find the people who are described by our statistics. It's there – not in the numbers – that we find cause. Statistics can document a pattern and tell us how consistent it is. Reporting tells us why it happens.

Sarah Cohen can be reached by e-mail at
sarah@nicar.org

Required reading

By Richard Galant
Newsday

From all the data you accumulate, from all the interviews you conduct, from all the documents you inspect, you should come up with one single, simple theme supported strongly by the data. Think of it as the skeleton on which everything else will hang or the classical music theme that will serve as the basis for all the variations. Everything in the story – the data, the interviews, the documents, the examples and anecdotes – will in some way relate to the theme, either by supporting it, expanding it, explaining its limitations or reacting to it.

Ensuring the story gets read

Write a lead that exemplifies the theme. If you're writing about a pattern of discrimination against minorities by a local hospital, don't begin with somebody who's the exception to the pattern. Instead, find the most interesting and dramatic case of discrimination you've got. The more you reduce the story to its essential elements, stripping away extraneous detail, the greater the impact on the readers. The more you begin with a story about people rather than numbers, the more likely you are to hook readers.

You should try to summarize your theme in a sentence or at most a lead paragraph. Think of it as the hard news lead you would write if you chose not to begin with an anecdote.

Keep it simple

You spent three weeks on the story, interviewed 50 people, examined 212 documents and put 6,000 records into a FoxPro database. You used a Pentium 200 with 64 megs of RAM to crunch the data. That's impressive and important, but why should the reader care? Don't let it get in the way of telling the story.

Story elements

Why do people read 500-page novels but complain they don't have time to read newspaper stories? Often, we try to give readers stories that are long on detail but short on compelling narrative that drives people through to the end. Narratives keep people reading by introducing believable characters and describing what happens to them over time. The best stories of this type have a beginning, a middle and an end. Reporting these is tough. You need to nail down the details of the characters, the color, the dialogue that you might not gather in reporting a more traditional survey story. There's no excuse for embellishing the details or creating "composite" characters. But if you do the reporting painstakingly and craft the story carefully, there can be a big dividend in more readership.

Avoid large, uniform sections – chunks of data or extended interview segments – in the story, if possible. To keep readers interested, create a "rocky road" mix of anecdotes, key findings, data and quotes.

Protect your findings

Be conservative in interpreting the data. Try to imagine all the potential arguments that could be made against your central theme. Bounce it off experts. You may have to go back to the data to make sure your findings and methodology are bulletproof. Give prominent play to the point of view of those who can make substantial arguments against your central finding.

"Killer comparisons"

Strive for ways to derive statistical findings that can be stated fairly in the most concrete terms. In *Newsday's* series on racial discrimination in mortgage lending, an analysis of nearly 100,000 mortgage loans found that African-American applicants are rejected nearly three times as often as whites are. That was stated prominently. But the story went on to derive another point from the same data set in a powerful way: "The gap is so wide that a black applicant making more than $150,000 a year stands a greater chance of being rejected than does a white person making less than $35,000 a year."

Home to papa

Bring it back home at the end of the story. Explain what ultimately happened to the people you wrote about at the top. The last few grafs of a story can be the most powerful and can leave readers with a lasting impression.

Richard Galant can be reached at (516) 843-3274 or by e-mail at rgalant@newsday.com

Search for news – not noise

By Nora Paul
Poynter Institute

"Hi all. We do not yet have access to Lexis/Nexis, DIALOG, etc. in our library yet but plan on (I am pushing for it) getting at least one of them soon. In the interim, however, I'd like to have some access to newspaper articles online (preferably free). I'm trying to avoid having to go to individual newspaper archives to search. General searches on search engines bring back some good results but often I have to weed through a lot of junk. Any suggestions from others out there who are in the same situation? Thanks."

Could you have written this message? Do you wish there was a way to get at the archives of newspapers or search their Web sites for articles easily? To help you accomplish that very thing, this column profiles.

Web sites with news search services or listings of news archives that enable to cast that net out to the 'Net and snag news stories only.

Small Hours / News Archives
www.aa.net/~rclark/archives.html

Rod Clarke, the creator of the list, includes weekly and monthly newspapers and the archives of news agencies and newsletters from organizations such as Amnesty International or U.N. Daily Highlights, and even some television station sites.

List Organization: Organized by regions of the U.S., Canada, Western Europe, Eastern Europe and Russia, Mideast Africa, South Asia, Asia, Oceania, Central and South America. The links take you to the homepage of the newspaper or to the search page for the archive.

News Hunt
www.newshunt.com

List Contents: News Hunt is a hybrid – part "scout page," part search engine facilitator. It gives you a listing of news publications with free searchable archives. Click on one of the 66 listings and you get a screen with a search box into which you can type your request.

List Organization: News Hunt is organized by categories of news needs: general archives arts & entertainment papers, breaking news,

business, popular, technology, and weather. Another section – newspapers by state – provides a listing of all the publications with archives in each state, creating a very useful compendium of both mainstream daily publications and niche and university publications.

Special Libraries Association / News Division Newspaper Archives on the Web
sunsite.unc.edu/slanews/internet/archives.html

This classic "scout page" was the brainchild of Margot Williams, metro desk researcher at the Washington Post. She was teaching an Internet course at George Washington University and turned her students on to the task of locating daily newspapers with available archives. The resultant list was augmented with additional links.

List organization: This long chart, expanding to include non-U.S. archives, includes links to the news sites' homepages, specific links to the archives within the site, database start-up date, and the cost of searching the site.

News Index
www.newsindex.com

"The most comprehensive, up to the minute News Search Engine on the planet." Begun in April 1996 by the entrepreneurial Sean Peck, this innovative service attempts to cull current articles from news sites from around the world.

Site contents: Including the major news sites and Reuters, it is an index of current articles only, not an archive. In fact, the service supposedly re-indexes the target sites every 1-2 hours. As such, it is a good way to get a cross section of contemporary coverage but it is *not* a good way to research old stories.

Other features: The custom search service delivers to your e-mail box articles fitting the profile you've selected. You can create up to five profile searches

News Works
www.newsworks.com

I don't know how long this nice little service from the ill-fated New Century Network will

Archive digs

be around, but use it until it dies a lamentable death. The feature most useful to journalists and others seeking content from a variety of news sources was the Search Service, which indexed and searched the contents of about 120 newspaper Web sites using Infoseek's *(www.infoseek.com)* spider technology.

Total News
www.totalnews.com

Soon after its launch, Total News got itself into hot water with news site publishers, who objected to the frames because they made the retrieved stories stay within the Total News frame. It looked, the publishers said, as if it were a Total News story.

According to a New York Law Review story in June, 1997, "...the media companies brought a litany of claims, alleging that the defendants' practice of framing constituted misappropriation, trademark dilution, trademark infringement, copyright infringement, false designations of origin, false representations and false advertising, unfair competition, deceptive acts and practices and tortious interference with contractual relations." Total News lost the lawsuit. This is actually an advantage to the searcher. Unlike News Works' search, which keeps the results screen within its page and makes it difficult to determine which site you are on, the whole page from the news site remains when using Total News.

Site contents: Total News' search robot scans about 1,200 news sites and updates the database three or four times a day.

Searching: The system uses a Web crawler called Paradigm News Search. Be sure to read the search help files – you'll see some interesting things, including the fact that one or two letter (or numeral) words are ignored in the search. (So forget using Total News to find reviews about the band U2's recent tour!)

NewsBot
www.newsbot.com
www.wired.com/newbot/

Using HotBot's search engine, this site is all about time. The search box lets you put in a search term and then select how current you want the items retrieved to be: 6 hours, 12 hours, 24 hours, 48 hours, 4 days, 7 days, whenever. You can also select certain types of news you want retrieved: all news, business, politics, entertainment, health, sports, technology, world. Each of the retrieved items has a date and time stamp on it in red to help the searcher select the most recent postings. The display of results is crisp and clean.

The bad news, though, is there is no clear listing of what sites are searched or how often. More bad news is the fact that the database is not purged of links to URLs no longer containing the news story. This is a big problem with Web-published news search databases. NewsBot, like its big brother HotBot, does provide some of the most innovative ways to focus searches through its search option boxes.

NewsTracker
nt.excite.com

Excite's news-only search service grants a number of ways to quickly scan coverage of particular events or specific features from a variety of news sources. Do a search and then have the results sorted by relevancy, with the lead of the story displayed. Select "Show Titles Only" for a listing of relevancy-ordered articles with only the headline and publication name displayed.

Click on "View by Publication" to see results grouped by publication and then sorted by relevancy. Or click on "View by Date" to get the results grouped chronologically first and then by relevancy.

YAHOO
headlines.yahoo.com/Full_Coverage

Get a quick list of top news stories and how they are covered by papers around the world. Click on one of the top stories they are following and you'll get a grouping of news stories by news, magazines & features, related Web sites, live coverage, or multimedia coverage. This is another of the clean, organized, reliable resources put together for 'Net citizens by the good people at Yahoo. Thank you, Yahoo.

But there is a bit of bad news, too. When you use the "Search" box to find stories in the news area of Yahoo, you get a listing of stories but no indication of where they come from.

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

Government felons

By Jeff Harris
WEWS-TV Cleveland

Jeff Harris attended NICAR's Basic Boot Camp in January 1998.

After a three-month investigation that combined a massive CAR project with old-fashioned, pen-in-hand research, reporter Ted Hart and WEWS-TV aired a two-part report in March that uncovered hundreds of convicted felons working in local government. Briefly, we reported:

• There were 375 convicted felons working as City of Cleveland employees. Another 90 individuals had been charged with felonies but eventually pled guilty to first-degree misdemeanors. One in every 20 city employees, we learned, was a felon.

• Of the 375 felons, 99 were convicted of violent felonies.

• Of the 28 city employees that served hard time in an Ohio prison, seven had claimed on their job application that they had never been convicted of a crime.

• Cuyahoga County has hired 151 convicted felons – 20 for violent crimes. An additional 17 were charged with felonies but pled guilty to misdemeanors.

• We also explored the quandary convicted felons face filling out job applications. We found that many are eager to find good, steady work but fear they won't be hired if they reveal their criminal background. We included interviews with the director of Community Re-Entry, a nonprofit group in Cleveland that assists convicted felons with finding jobs, and an ex-con who had found work through the group and put his life back together.

• Many local and state government entities have developed policy initiatives that encourage the hiring of convicted felons and work with groups like Community Re-Entry. This was not the case with the hiring practices in Cleveland. No initiatives were in place to recruit or hire ex-offenders.

Brainstorming

Ohio's prison system, we learned, maintained a database on every individual incarcerated in an Ohio reformatory. It included the name, date of birth, date of admission, term of sentencing, name of prison, county of conviction, description of crime(s) and projected release date of the prisoner. After $200 and a few days of tinkering and retooling by the brilliant NICARians in Columbia, we had our grubby hands on the database.

As we began to hash out our story, we addressed a number of issues and research strategy questions. For those felons on the workforce, is there a correlation between their crime and their position? Did these ex-offenders lie on their applications? What are the total numbers and how do they compare from city to city? Is there a pattern? What is a felon to do? Are there instances of convicted felons committing crimes after having been hired? Are there policies in place regarding the hiring of felons? What are the application policies on criminal background checks.

Hunting and gathering

We decided to broaden the scope of our investigation beyond Cleveland employees. In late 1997, WEWS began submitting public record requests for names of employees, job titles, departments, dates of hire, and dates of birth from a broad range of local government agencies. For the most part, the cities turned the FOIA requests around within reason and at reasonable costs. I did encounter the common line of defense employed by bureaucrats: "That's not public record. We can't do that. Well, we may have to do it, but we don't have to do it on disk. You'll have it by 5 o'clock."

Realizing the vast majority of convicted felons never see the inside of a prison, we concluded our developing story would be most thorough and accurate if we were able to cross employee databases with county clerk of court records – in which we would pick up those sentenced to county jail time and probation.

Thus began the tale of two county court databases. WEWS submitted written requests to the clerks of Cuyahoga County, Ohio's largest, and Summit County, nearly one third the size. We sought the name, date of birth, case number, criminal charge and disposition of every criminal defendant in their courts. Within three days, Summit County e-mailed a database containing the exact in-

Recurring debt

formation we asked for – all for a total of $86.

Cuyahoga County handled the FOIA much differently. A month after our FOIA request, we were informed it could be fulfilled for $10,000. News director Jim LeMay pounced on the phone and called a county commissioner. In his tactical tirade, he used words such as "gouge" and "outrageous" and "lawyers" and "court costs" and "in the public interest." The next day the price dropped to $4,000. And if we would settle for just name, date of birth and case number, the database would be free. We settled for the free one. Three months after our initial request, the county delivered the database to WEWS on CD-ROM.

Confounding fields

The major stumbling block in getting the information into one database was that every city recorded employee names differently in their tables. Most complicated were tables that included middle initials and Jr./Sr. designations. I used Excel's Text to Column command to parse the confounding fields, often pasting only the last name column back into the original employee spreadsheet. I knew my Access cross match of the court and prison databases would entail last names and dates of birth, so the last name column would be all that I needed.

Last names and dates of birth are constants. First name matches leave out a large chunk of true matches because nicknames or initials may appear in one table while an "official" name appears in another. They may appear in a first name field or a field of their own. And first names are misspelled more often than last names.

Matching first names by eye allows for the perfect integrity check and for the opportunity to truly examine the results of the query record by record.

Weeding through several hours of queries and the resulting tables, we found more than 700 separate last name/date of birth matches, with first name "eyeball matches" totaling more than 1,000 separate cases.

Reality check

Our remaining task was confirming that the employee and the ex-con were one. The county tables contained 700 individuals who had been indicted in felony court over the past 30 years, but we had no idea whether these people had felony convictions, and if so, when, and for what crimes.

To confirm convictions, we took our list to the public computer terminal at the clerk of court and punched in case numbers. Initially, we selected the case numbers of employees who worked in positions with a high rate of contact with either the public or money, or who were involved in public safety.

Our goal was to find potential subjects to highlight for our story. Our first sweep yielded a host of compelling cases: a deputy treasurer convicted for sexual battery; a recreation center director with drug trafficking convictions; and a park maintenance supervisor convicted of involuntary manslaughter for his part in a drug store robbery and murder.

Unfortunately, it took us weeks to narrow our list of subjects, find them, and then get interviews and pictures. One week before airing the story, we didn't have final numbers. We should have completed confirming charges on the entire list and compiled our final numbers before finding the subjects for our story.

Our news desk received dozens of calls after our story led the 11 p.m. newscast on March 2. All were negative. Viewers thought we were picking on people who had paid their debt to society. Many complained that we had needlessly ruined the lives of people who had reformed. After part two ran the next night – discussing the difficult employment problems faced by ex-offenders – we received calls asking just which side we were on.

The three employees we highlighted as having lied about their felony convictions were suspended by the city pending termination. After requesting our list of 375 felons, the city has opened its own investigation. They will review the job application and convictions in each case to determine whether disciplinary action is warranted. Cuyahoga County has also opened an investigation and fired the deputy treasurer convicted of sexual battery.

Jeff Harris can be reached at (216) 431-4231 or by e-mail at harris@wews.com

Bits, Bytes and Barks

Online CAR Projects

To view a listing of links of recent computer-assisted reporting stories posted on the Web, point your Internet browser to *www.ire.org/resources/conferences/training/carprojects.html.* The site includes a description of the stories as well as links to IRE award winners. If you would like to see a story added to the list of links, send Jack Dolan an e-mail at jack@nicar.org

Upcoming IRE & NICAR Events

Besides the bustle of investigative journalists discussing their craft in the Big Easy this June, there's a slew of other IRE and NICAR events coming soon, including on-the-road training, seminars, regional conferences and boot camps. For a full listing, see the IRE Web site at *www.ire.org/calendar.html.*

Database Library Additions

NICAR has recently acquired updates to several of its databases: the U.S. Coast Guard's boat registration database, the Immigration & Naturalization Service's database of naturalized citizens, the Government Services Administration's database of federal procurement contracts, and the Bureau of Alcohol, Tobacco & Firearms' database of gun dealer licenses.

For a description of available databases or a downloadable order form, see the Web site at *www.nicar.org/data.*

Campaign Finance

The Campaign Finance Information Center continues its pursuit of acquiring campaign contribution data from as many states as possible to allow national contribution patterns to be discerned.

Downloadable data from ten states--Florida, Idaho, Illinois, Indiana, Kansas, Kentucky, Michigan, Minnesota, Ohio and Wisconsin-- has been added to the CFIC Web site, located at *www.campaignfinance.org.* Comparisons can be made by joining tables between these state files. CFIC hopes to enable such comparisons to be made online once common fields are established and standardized.

Also available at the Web site are links to online databases maintained by various states themselves, though not all can be downloaded.

Direct your browser to *campaignfinance.org/cficmap.html.*

CFIC Mailing List

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