June 1997

Uplink

A newsletter for the National Institute for Computer-Assisted Reporting

Summer Session

Uplink Update

This month's Uplink offers a smorgasbord of CAR offerings to get your summer started off right.

Michael J. Himowitz tells how *The Baltimore Sun* discovered "dreams, debts and demolitions" in the city's public housing.

Tom Gaumer from *The Cleveland Plain Dealer* tells how he investigated the FDA's experiment procedures and found alarming results.

Ames Alexander of *The Charlotte Observer* teaches the ABCs of day-care investigations.

Sarah Cohen uses her statistics column to share how the *Arkansas Democrat-Gazette* tackled the tricky, but important job of ranking area schools.

Bob Sullivan of MSNBC also tells about his first computer-assisted reporting project, which uncovered bid rigging at the FCC.

And Nora Paul gives the must haves of Internet sites in her monthly Internet column.

Inside

Housing Investigation

Dreams deferred

By Michael J. Himowitz
The Baltimore Sun

Martha Fields' West Baltimore rowhouse was a dump with a leaky roof that brought down her bedroom ceiling and forced her into the street.

Barely subsisting on Social Security, the 65-year-old woman had no money to fix up the eyesore, and when neighbors complained, the city responded with a work crew. Fields thought that was wonderful — she was getting help at last.

What she wound up with was a boarded-up derelict saddled with $37,000 in liens that she can never hope to repay. In fact, with interest accruing at 20 percent a year, Fields now owes three times what the house is worth.

But no one is ever likely to buy it.

Her tale of woe and others sparked a year-long investigation of Baltimore's decaying housing stock by *Sun* reporters John O'Donnell and Jim Haner. The outcome was a three-part series titled "Dreams, Debts, Demolitions" that ran April 6-8.

The big picture

With the help of a sophisticated computer-assisted reporting effort, Haner and O'Donnell discovered that Fields' case was far from unique.

They found a well-intentioned government effort gone awry, a campaign to save neighborhoods that actually encouraged their abandonment while lining the pockets of politically-connected contractors.

The result: a city where 10 percent of the residential addresses are owned or controlled by the government. A city with as many as 40,000 vacant houses and empty lots. A city where millions have been spent repairing and renovating homes that were torn down within a few years.

O'Donnell and Haner outlined the decline of a politically popular program set up in the 1980s to battle slumlords. It gave the city the power to repair rundown houses and bill the owners for the

Continued on page six

CAR & FDA

Experimenting with Drugs

By Tom Gaumer
Cleveland Plain Dealer

When the Food and Drug Administration inspects drug experiments, it finds more than half the time that the human subjects have not been sufficiently informed about their possible risks and benefits.

That's the conclusion in articles by *Plain Dealer* reporters Bill Sloat and Keith Epstein. The series ran over four days beginning Dec. 15.

The project began more than a year ago after Sloat and Epstein finished stories about Norplant and how some women did not give writ-

Continued on page two

Uplink
June 1997
Volume 9, Number 6
A newsletter for the National Institute for Computer-Assisted Reporting

Uplink is published every month by the National Institute for Computer-Assisted Reporting, 138 Neff Hall Annex Columbia, MO 65211. (573) 882-0684. Subscription price is $40 for IRE members, $60 for nonmembers. Postmaster: Please send address changes to NICAR. Send e-mail to jeanine@nicar.org

NICAR is a joint effort of Investigative Reporters and Editors and the University of Missouri School of Journalism. NICAR services include hands-on newsroom training on computer-assisted reporting, special academic and advanced training in data analysis. NICAR is supported by grants from The Freedom Forum and other foundations intended to help the institute deliver its services nationwide to news organizations and associations.

Continued from page one:

FDA investigation

ten consent when Norplant was tested on them. Managing Editor Gary Clark suggested that the two try to find out if there were other instances where the FDA found inadequate consent.

The two reporters obtained more than 5,000 pages of federal documents along with a computer database containing the FDA inspections from 1974 through 1995 of clinical research involving experimental drugs being tested on humans.

The reporters found the database almost by accident.

They had filed many FOIA requests for FDA inspection reports, but didn't get any on the grounds there were none.

They started calling employees from the FDA and someone told them to ask for the 483s, which was the actual inspection report. Armed with that information, they got a printout of some 483s.

But then someone mailed them a huge printout of the actual database. With that, they knew for sure the database existed on a computer.

It took another three months, but Epstein, who works in *The Plain Dealer's* Washington bureau, finally got the records on floppy disks.

That database provided numbers that were the basis for much of the series as well as most of the information for the graphics that ran with the project.

The database contained 8,039 records. There was no single field for each inspection so we used FoxPro 2.6 for Windows to calculate that the FDA conducted 4,154 inspections of medical researchers.

The database had a column for type of deficiency, and we found that in more than half the inspections — 2,201 — the consent was inadequate.

The database even revealed 47 instances where there was no consent.

One of the major facts Sloat and Epstein wanted to know was how federal research locations compared with those at universities, hospitals or private facilities such as clinics.

But the database had no classification indicating what type of facility was being inspected.

In addition, it had the usual pattern of using different spellings or words to describe the same institution.

So Sloat and I spent a week in early 1996 going through the entire database (we were able to throw out some records that did not have a location).

We added a new column and gave each location a code letter such as G for government or U for university. Private facilities or those we could not identify went into the same category, which we called other. But we were able to identify most of them from the name.

Once that new database was completed, we were able to go far beyond the information available at first.

We found out very quickly that the government locations, such as Veteran's Administration hospitals, had the worst track record for providing information to the people used in experiments.

Because we now had a unique number for each location, we were able to pinpoint the government locations and universities that had the most violations over the years. All these facts were used in graphics that ran every day.

To make the story interesting, Sloat and Epstein had to find the names of people who were part of studies. The task was painstaking because identities of patients were omitted or deleted from many records. In other cases, the government used code names.

Sloat and Epstein traveled extensively throughout the country and even to Europe tracking experiments inspected by the FDA.

Sloat and Epstein even found people who died during a drug experiment, including a Pennsylvania woman who received a substance also used as a gasoline additive to try to dissolve gallstones.

That woman was only known in records as TE3. Sloat and Epstein were able to match public death records with FDA documents to find the clues that enabled them to identify the woman as Laura Michalski.

They talked to Michalski's widower and children who were unaware that Michalski had been part of a drug experiment.

Tom Gaumer can be reached at
(216) 999-4500, or send e-mail to
TGAUMER@plaind.com

Day-care discovery

Beyond the nightmares

By Ames Alexander
The Charlotte Observer

We'd all read stories about day-care horrors, the tales of abuse and neglect that leave readers with a sense of disgust and fear but little understanding.

So when other *Charlotte Observer* staffers and I decided to assess the state of child care in North Carolina, we knew we needed to do more than simply lay out the parental nightmares.

We needed an authoritative way to find out how often day-care kids are mistreated, where it's most likely to happen and why.

The challenge was daunting, but computer databases made it less so.

Using Foxpro, we analyzed tens of thousands of records in databases kept by state day-care regulators. That — along with scores of interviews, visits to day cares and trips to state filing cabinets — helped us illustrate the consequences of the state's lax day-care rules.

Here's some of what we found:

Each month, more than 100 North Carolina children are hurt so badly in day care that they require medical treatment. That's partly because day-care staffing requirements in North Carolina are among the nation's most lenient.

Children are more likely to be neglected or abused at centers that meet only minimal state requirements. Complaint rates are particularly high at the state's smallest day cares, which are based in homes and subject to the least state scrutiny of all.

The rate of substantiated complaints at the state's two largest chains — KinderCare and La Petite Academy — is more than twice that of the average center in North Carolina.

Over the past five years, state officials have discovered nearly 1,700 North Carolina day cares operating illegally. More children were neglected and abused in those homes than in all the legally registered child-care homes combined.

Inspectors don't punish or close many problem day cares. Though state inspectors substantiated more than 1,400 complaints against day cares during the past two years, they revoked just 36 licenses and issued only 138 fines.

The databases proved useful in other ways as well. They steered us to the child-care centers with the worst track records. And they helped us produce comprehensive summaries of all recent state actions against day cares in the Piedmont. Lots of work — and a little resourcefulness — went into those summaries.

We used the databases to generate a list of all recent state actions against day cares in the 10-county region. Trouble was, the computer files didn't give us a full account of what happened and how the day care responded to the state's action. So we took our list toRaleigh, and pulled the files on each day care cited.

Armed with laptops, we plowed through the files and created our own databases, entering information about what happened and how the day cares responded.

To be fair and accurate, we later tried to contact the owners of each day care. Then we printed the summaries in their entirety.

If you're looking for databases with information on child-care centers, start by asking the state agency that regulates them. In North Carolina, it's the state Division of Child Development, an arm of the state Department of Human Services.

Without too much coaxing, we got databases with a great deal of valuable information: complaints against day cares; injuries among children; names of day care owners and their phone numbers; whether each day care was licensed or registered; how many children each center was licensed to care for, etc.

One thing to keep in mind: It's not fair to compare the raw number of complaints at each center.

You need to calculate complaint rates, based on the number of children each center is licensed to care for. And when figuring those rates, it's best to use substantiated complaints, rather than total complaints. Not every parent's gripe is well-founded, experts note.

In the end, computers could do only a small part of the work. We spent much of our time visiting day cares, poring through state records and interviewing teachers, parents and experts.

But used well, databases can make your child-care stories more powerful, accurate and downright interesting.

Ames Alexander can be reached at (800) 283-6459, or send e-mail to amesalex@charlotte.infi.net

Several of the stories featured in this issue of Uplink including the *Cleveland Plain Dealer's* FDA story and the *Arkansas Democrat-Gazette's* school ranking story are available through the IRE Resource Center. To order copies of the stories and tips from the reporters who did them, call (573) 882-3364.

Stats from the road

Head of the Class

If you have any ideas, examples or questions about statistics, e-mail Sarah Cohen, NICAR's On The Road trainer, at sarah@nicar.org. Her column appears every month and looks at useful statistical tools or other number techniques that can help boil a story down to its core elements.

By Sarah Cohen
NICAR

When the *Arkansas Democrat-Gazette* set out to measure the quality of its schools, it didn't look to fancy statistical techniques to do its work.

Instead, reporters and editors thought through the project, read about school reports, evaluated the information they could get and culled the reporting of three experienced education reporters.

Then they came up with a simple set of criteria that they would use to define the word "best" for their readers.

Yes, test scores are good measures of a school's success.

But when parents look at schools, they often look also at the environment, diversity and other measures of quality.

Democrat-Gazette reporters took that into account when devising their index of excellence in schools.

They included measures of teacher training and coverage, racial disparity, performance, parent activity, and motivation. And they combined them without the regression techniques that have been used elsewhere to find over- and under-performing schools based on poverty.

"We may do that in the future," said Jeff Porter, the computer-assisted editor who crunched the numbers for the series. "But what was interesting was that the elementary and junior high schools that were at the head of the rankings weren't in the wealthy parts of town. The elementary school at the top is almost a rural school, on some road that kind of meanders around west of town."

Beginning at the beginning

The effort shows why thinking through a project as perilous as the annual school reports is just as important as learning advanced statistical techniques.

Educators tell us there's no way we'll compare schools fairly. But if our job is to provide our readers and viewers with information, not hide it, then our other job is to do it as responsibly as possible.

That's where *Democrat-Gazette*'s indexing technique comes in. In Excel, Porter essentially ranked each indicator, then added up the ranks. The school with the biggest total came out on top.

The only numbers trick he played was to turn the rankings upside down: The school with the "best" performance got the lowest rank, giving it the most points toward the total score.

The school with the "worst" performance got the highest rank — the No. 1 we always think of in rankings — contributing the least to the total score.

Combining upside-down rankings of many indicators is one way reporters can cut the school reports down to size.

The approach seems embarrassingly simple for those who have struggled with the numbers morass of school reports.

Lots of indexes are done this way.

Most business "Top 100" lists are driven by these kinds of rankings. So are the "best places to live" indexes.

Even *U.S. News & World Report*, which ranks institutions from hospitals to colleges, uses a system something like this.

It might be more sophisticated, but the basic idea is the same — figure out what is important, assign a score for each element, and come up with a total that puts everyone on the same footing.

Is an index made up of rankings scientific?

Not entirely.

Does it have drawbacks?

Certainly.

You've lost a lot of information that you'd retain if you used z-scores instead.

All of these indexes, though, use some combination of judgment, reporting and sometimes even polling involved in deciding what's important.

Some words of caution:

• Figure out your indicators carefully.

If you decide that racial disparity is a bad thing for schools, then you have to find some indicator of this that will treat all schools equally. *Democrat-Gazette* reporters had criteria laid out by the court system to decide how to measure it.

The lesson is clear – find some reputable and impartial source for your standards whenever possible.

Continued on page five

Continued from page four:

Grading the schools

Adjust most of your numbers for the number of students. This is the "per capita" calculation that NICAR classes harp on so often.

Arkansas used many measures "per 100 students" instead. It's the same thing.

• Make adjustments that educators say matter, but only when they do.

Many school report stories use some kind of adjustment for poverty, often using the percent of the students who qualify for free or reduced-price school lunches.

You might want to give schools with high poverty rates (or other difficulties like English language skills or high migration rates) a leg up in your ratings.

By focusing on more than just test scores as a measure of quality, *Democrat-Gazette* reporters and editors didn't need to make these kinds of statistical adjustments.

• Throw out redundant indicators or others that don't tell you much.

School systems often throw out reams of information, much of which is useless in judging the quality of schools.

For instance, New York's Web site scores of indicators, some of which are nothing more than the percentage of kids who passed competency tests in each subject for each grade. This doesn't help when they're always 99 or 100 percent.

• Decide what you'll do if there's a lot of missing data.

In Arkansas, like many other places, reporters simply couldn't get reliable information for some grades.

You may decide to focus only on the very early grades, then junior high schools.

You may have to skip a story on middle schools if there aren't enough to make a fair ranking.

Or you may decide to include an average of the schools' other indicators as an estimate. Just don't let the missing values sit there, though. It will pull down the scores of otherwise fine schools.

• Collect raw numbers, not rates.

Democrat-Gazette reporters wanted to make sure each indicator was calculated the same way.

So instead of collecting the usual drop-out rates, reporters had to ask school officials for the raw number of dropouts.

This is different, and raises the ire of some school officials, Porter said. But it's sometimes the only way to make sure your numbers are all calculated the same way.

Once you've gotten your indicators, though, making an index of rankings is straightforward. Get your final indicators into Excel, in any order.

Now use the RANK() function to assign the scores for each indicator. Then add up all the scores. You have to figure out which direction is "good" and which direction is "bad" for each indicator.

For instance, a lower teacher-student ratio is usually considered better than a higher one. So a low number is "good."

Or a higher test score is better than a lower one, so a high number is "good."

Here's how a formula might look for a "high is good" indicator.

The RANK() function requires three pieces of information: the number you want ranked, or this test score; the group you want it ranked within, or the entire range of test scores; and the direction you want it ranked.

A "1" means that a higher number will end up higher.

A "0" means that it will rank normally, with a lower number assigned to a high score.

```
=RANK(B3,B$3:B$77,1)
```

Note the dollar signs.

They mean that, as you copy down, the formula will remain correct. But they also mean you can copy it across, only changing the "1" at the end when you come to an indicator in which "low is good."

Of course, using more sophisticated techniques can help you find the stories in your school system.

In the July issue of Uplink, we'll use the statistics column to profile a set of stories that used statistical techniques to find schools that overcome difficulties well, or sit on their advantages.

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

To see the *Arkansas Democrat-Gazette's* full story about the quality of Arkansas schools, go to www.ardemgaz.com

Continued from page one:

Demolishing dreams

The complete housing series can be found on SunSpot, *The Baltimore Sun's* Web site. Point your browser to www.sunspot.net/news/special/housing/housing1.html

cost. To guarantee a return if the owners didn't pay up, the city put liens on the houses that would be satisfied when the properties were sold.

But the system wasn't designed to deal with a 20-year exodus of 200,000 people that sent inner-city property values plummeting.

Instead, the liens encouraged thousands of landlords to abandon their property and turned low-income homeowners into debtors.

Thousands of homes became albatrosses burdened with liens so large that no one could afford to buy and renovate them.

With interest accruing every year at punitive rates, some properties worth $10,000 have been saddled with liens worth $100,000.

"We found many cases of vacant houses that were definitely worth rehabbing if it was affordable," O'Donnell said, "but it isn't affordable with a $30,000 lien on the property."

The *Sun's* reporters also found a housing department so backlogged and disorganized that it has no idea how many properties it actually owns.

They uncovered an aggressive but willy-nilly program of demolition that has left some crumbling blocks untouched and others gap-toothed.

And they also found a bureaucracy churning along with an outdated system that routinely mails out bills for repair work to long-dead homeowners and long-departed landlords.

But within days after the housing series ran, Mayor Kurt L. Schmoke announced a drastic slowdown in the city's demolition program and promised a new look at the liens and interest payments that have tied up so many properties.

In the beginning . . .

"Dreams, Debts, Demolitions" had its genesis in complaints from Fields and other angry residents who called Haner after a 1996 series revealing that Baltimore housing officials owned neglected inner-city properties (The series was a finalist for the 1996 Pulitzer Prize).

Like most effective CAR projects, "Dreams, Debts, Demolitions" combined old-fashioned legwork with computer analysis. The computer pointed the way, but Haner and O'Donnell dug out the story by talking to people and pulling original documents.

The *Sun's* electronic news department used a variety of housing-related databases and other records, mostly purchased from the city.

Although we had plenty of anecdotal evidence that something was wrong, we used the computer to document the extent of the blight.

"The beginning of the trail was the lien database," O'Donnell said. "As we learned more about the process, we began using all the databases to develop information."

That meant identifying vacant houses and lots, city-owned residential properties and tax sale derelicts. Then we tried to figure out how the worst properties got that way, how much debt they'd accumulated, and who made money along the way.

Piece by piece, we built an electronic "paper trail" of decay by sifting through hundreds of thousands of ownership records, housing violation notices, liens and building permits.

Eventually, we were able to trace the history of thousands of crumbling homes. By linking the property tax and building permits databases, we were able to track the work done on every house, from repair through demolition.

At various times, we used the following databases: property tax assessments, city property liens, building permits (including demolitions), housing violation notices, housing authority payments, the city's vacant house database, city employees, campaign contributions, tax sale records, voter registration records.

Our primary tool was Microsoft Access, with some help from Excel. We also used Caliper's Maptitude (a thematic mapping program) to geocode the addresses of rundown properties and develop a map showing the worst housing blight.

Our most important assets were the Baltimore tax assessment database, which provided information about the owners and properties; the building permit file, which let us track city-sponsored repairs and demolitions; and the city's database of tax sale liens, which put us on the trail of houses that had been swallowed by the system.

We developed an inventory of city-owned residential properties by searching the assessment ownership field for city agencies.

We determined which were residential by looking for dwelling units and clues such as lot size.

We were able to analyze ownership in blighted areas quickly by displaying entire blocks of houses, looking for patterns.

To identify vacant residential lots, we first

Continued on page seven

More databases; Volunteers

By Andy Lehren and Cory de Vera
NICAR/IRE

The National Institute for Computer-Assisted Reporting continues adding and updating databases for journalists, and is enlarging its data analysis services.

Among the new databases are: the Centers for Disease Control's reports on AIDS patients across the country, the Food and Drug Administration's collection of complaints about drugs and medical devices, and the U.S. Census Bureau's Consolidated Federal Funds Reports, which allow you to track federal spending in your community.

The updated databases include: the U.S. Department of Transportation's reports of accidents involving hazardous chemicals, the U.S. Small Business Administration's 7A loan guarantees for companies, FBI crime reports, railroad accidents, federal contracts, and files on legal immigrants to the United States.

To continue enlarging its offerings, NICAR has numerous FOIA letters and other pending requests for data. We will announce those offerings in this newsletter and on the NICAR-L listserv as they are completed.

NICAR has been expanding its analysis services for broadcasters and newspaper reporters. This goes far beyond helping reporters process tricky tapes — from restaurant inspections to property records. NICAR is doing more and more data analysis for reporters. This year, we've begun analyzing records using mapping programs.

For more information about how NICAR can help you, call NICAR at (573) 882-0684 or e-mail Andy Lehren at andy@nicar.org.

We Want You

IRE is reviving and restructuring our system of state and regional coordinators and volunteers. If you'd like to volunteer to be a local coordinator or just work on specific projects, now is the time to let us know.

As a local coordinator, you might organize an evening reception for journalists in your area, or you might report back to IRE about the great investigations and CAR projects being conducted by local news organizations. We especially need people who will help us recruit new IRE members. Becoming a local coordinator or volunteer is a great way to get in touch with other journalists and keep yourself current with the best work in the business.

Strengthening our presence locally is an effective, grassroots way we can reach out to journalists who can't always make it to a regional or national conference. We know that IRE has a good thing going, and we need you to help get the word out.

If you believe in the IRE mission and have some time to lend, call Cory de Vera at (573) 882-2042 for a volunteer application form. We'll let you know who else has volunteered in your area, and send you a volunteer's information pack.

***The Baltimore Sun's* '96 series on housing was a Pulitzer finalist. To find a list of all the Pulitzer winners and finalists, many of whom used computer-assisted reporting, go to www.pulitzer.org**

Continued from page six:

Public housing crisis

searched for properties where the value of improvements was zero.

Then we filtered out properties without regular street addresses, which eliminated parks, alleys and rights-of-way. Finally, we eliminated lots too small or too large for individual houses.

To locate properties made irredeemable by the system, we matched assessment records against the lien database and flagged properties with liens that were way out of proportion to assessed value.

We were then able to develop a list of housing code violations and building permits for each of those homes. In many cases we found that the city had spent thousands to repair properties and then had them torn down.

Using the permits database, we developed a list of contractors who got city jobs to fix up and/or demolish old buildings. As the years went by, an increasing number of contracts went to firms whose owners had ties to city officials.

Work on the project continues. A few weeks after the series appeared, Haner and O'Donnell wrote that the city had spent $500,000 on a new crane to speed up demolitions. But the crane stood idle for a year because the only two employees qualified to operate it had taken a buyout.

The city couldn't pay enough to find replacements.

Michael J. Himowitz can be reached at (410) 332-6638, or send e-mail to mikeh@clark.net

Rigging the Results

Bob Sullivan's story on the FCC and supplemental charts can be viewed http://msnbc.com/news/73393.asp

By Bob Sullivan
MSNBC

Bob Sullivan graduated from the Missouri School of Journalismin May 1996.

MSNBC's investigation into bid-rigging in FCC spectrum auctions was an interesting mix of a long-term CAR project and spot CAR reporting.

My editor, Mark Pawlosky, had charged me with finding databases regarding the FCC's auctioning of spectrum bands to telecommunications companies.

There is so much money changing hands in these auctions — $23 billion has been raised since the auctions started three years ago — that the process was worth taking a look at.

With our telecommunications reporter, David Bowermaster, I had been in touch with Andy Lehren at NICAR regarding what data could be found or FOIA'ed.

We had asked for some pretty thorough data, not knowing what we were looking for, and NICAR was assembling it in Missouri.

A public fight

Then on May 1, Reuters reported that two small companies involved in the January auction for Personal Communications Services licenses (PCS devices) were in a public squabble.

High Plains Wireless filed a complaint with the FCC accusing Mercury PCS of "bid-signaling" in January's DEF block auctions.

The complaint spurred an investigation by the Justice Department, and reports said the last three digits of bid amounts were somehow involved.

That was the burst of sunlight which we needed.

Now we had something to look for. Companies issued bids that ended in unexpected numbers, like $1,100,136, as a signal to competitors.

The signal might tell others to stop bidding on that market, or it might contain a message about bidding in another of the 1,479 licenses available in that auction.

In any case, the hypothesis was that companies used this signaling to limit competition for licenses and by mutual unspoken agreements, keep prices down.

Now, NICAR was able to point me to the enormous and unruly data the FCC had posted on its Web site regarding the PCS license auctions.

In the DEF blocks, there were 276 rounds of bidding. Sure, all the data was on the Web site — in over 1,000 different .dbf files.

There were six files for each round — one with all the round's bids, one with the high bid before the round, one with the high bid after the round, one with that round's withdrawn bids, etc.

A sample test

At first, we agreed it might be futile to download all of it. But we figured there might be a way to test a sample of the data to see if it might bear fruit.

First, we looked at the winning bids for the 1,479 licenses, to see if anything about the last three digits of the winning bids stuck out. Not really.

So then we cherry-picked every 20th round, and ran some tests on a mini-version of the data. We had some notion that bidders were allowed to lower their bids as the process progressed if they weren't challenged in a market.

That seemed odd, so we went hunting for markets where the winning bid was less than a bid made during the rounds. We found a few, but that was also basically a dead end.

It did point to a few markets, though, and when examined, the bidding in those markets exhibited what seemed to be odd bidding behavior — and about half of them involved a company named Western Wireless.

There seemed to be multiple instances of the same dollar amount being bid by the same company for the same license. That didn't make any sense — why would they have to make redundant bids?

Expanding the search

We now figured it was worth our while to get a picture of what happened in all 276 rounds of bidding.

It took most of a weekend, but we grabbed all the bids made in all 276 rounds. We then imported the tables into Microsoft Access.

Had I more experience, I would have written some kind of script to bundle them all into one table.

But not knowing any other way, I used one

Continued on page nine

Continued from page eight:

Bid rigging revealed

append query at a time to compile into one table all the bids made in all the rounds on the DEF block.

Fortunately, each round had a field with the round number, so we could assemble them into one table without losing data or being forced to add markers.

The results were now quite good – I built a quick front end which allowed me to query one market at a time and showed the bidders and bids as they marched from first round to winning bid.

The results and more work

What emerged? The most noticeable was Western, which seemed to have a habit of making the same bid over and over — in fact, it won 10 licenses in the 136th round with that kind of redundant bid.

There were plenty of other signals, too — like Comcast's bid for a Philadelphia license that ended in 346, when that was also the city code for Philadelphia — or Western's three round 83 winning bids, which ended in 083.

But we only had half the picture, and Bowermaster's reporting here pointed us to the other half.

Consultants told him that a frequent method for signaling was to make a bid, withdraw it quickly (sometimes in the same round), then make the same bid again.

So, it was back to the Web site for the 276 .dbf files of withdrawals.

I then had to merge the withdrawals and the bids into one table — because my goal was to follow all relevant activity in a market as bids progressed.

To do this I had to add a field, because the data didn't indicate whether a bid was a withdrawn.

I just added a withdrawal field to my combined table, then pasted in the withdrawal tables, with "default for new records" set at "yes."

The 552 tables assembled into one, I now had a complete picture of the competitive bidding for a market.

Sure enough, those 10 licenses Western had won in Round 136? In round 135, they had withdrawn a bid for the same amount in every case.

When we called Western to comment on our findings, they refused, but did say they had been contacted by the Justice Department.

But there was no reason to pick on Western — dozens of other companies had done similar things.

Omnipoint Corp., bidding as OPCSE-Galloway Consortium, made and then withdrew bids on 171 occasions — 65 times Omnipoint entered bids and then withdrew them in the same round.

On our Web site, we were able to post round-by-round bids for about 20 communities which showed this kind of bid-signaling.

Bowermaster's reporting, combined with work from Brock Meeks, revealed that this behavior actually had its roots in "game theory," which was applied in the design of the auctions.

Game theory is elaborate, but suggests temporary alliances (as you might do in the board game "Risk") and poker tactics to succeed in battle.

Surprise — some of the consultants who designed the auctions actually consulted with companies doing the bidding.

Several say the bid-signaling was fairly well-known and accepted as part of the process — but not by all.

The Justice Department is currently trying to determine if bid-signaling was in fact bid rigging — if it was fair game or collusion.

Mercury PCS has replied to the FCC saying they did use these tactics but argued that they were not illegal, and used our story extensively in their request for dismissal of High Plains' complaint.

The debate on legality continues — as does the debate on whether "everybody was doing it."

After all, even if it wasn't illegal, and most companies knew about it, those who didn't were disadvantaged.

And the FCC did issue a warning after round 135 complaining that withdrawals were being misused. The Justice Department investigation is ongoing.

The good news is, this is a CAR story that broke off a beat, and because we were already doing some homework in the area, we pulled together a pretty exhaustive story five days after the initial story broke.

Bob Sullivan can be reached at
(202) 936-7751, or send e-mail to
Bob.Sullivan@MSNBC.COM

Looking to take your first venture into computer-assisted reporting? Let Brant Houston help. To order his "Computer-assisted Reporting: A Practical Guide," call (573) 882-2042. Cost for the book is $26 plus shipping.

Tech Tip

What's Your Function?

If you need more CAR assistance, you can order audio tapes of presentations given at either the NICAR or IRE national conferences. To order a tape or to get a list of the offerings, call Gaylor MultiMedia Inc. at (615) 361-3611.

By Richard Mullins
NICAR/Missouri School of Journalism

The FoxPro programming language is composed of command and function words. You use these for simple interactive commands and for programs, simple and complex. It's also possible to combine the existing functions and command words, package and label them as a new function, then use that new function just like the built-in functions.

This make-your-own-tool capability is a fundamental feature of most programming languages. I'll come back to the uses and advantages of making functions. First, let's make a really simple one.

The first function we'll write doesn't really add anything new to FoxPro, though that's the usual reason for writing one. The function noDot( ) returns a string with all the periods stripped out. It's equivalent to, but shorter than the second command.

•The function we'll write will be used like this:

```
? noDot(address)
```

• It is equivalent to:

```
? strtran(address, '.' , '' )
```

• Example code

```
FUNCTION noDot
Parameters cString
RETURN strtran ( cString, '.' )
```

That's it. Type three lines, save it, and you have a new function for FoxPro. The manual calls them UDFs, or User Defined Functions.

The next function we'll write doesn't add anything new to FoxPro either, but it's a handy substitute for its hard-to-memorize-and-type equivalent.

The function rSplit( ) returns the part of a string that's on the right side of a separating character.

• The function we'll write will be used like this:

```
? rSplit(name)
```

• It is equivalent to:

```
? substr(name, at(',', name))
```

This one is a bit more complicated. Also, we'll write it in several passes, adding and explaining features, improvements along the way.

• First version

```
FUNCTION rSplit
   parameters cString
   nSplitPos = at (',' , cString)
RETURN alltrim(substr (cString,;
  nSplitPos + 1 ))
```

Most of the time, the comma is the separating character in the string we want to split. Our function, as written, assumes that. However, if we want to use this function to split apart a string based on another character, we'd have to edit the code to change the quoted character in the second line, where the at( ) function finds its location.

When we were done with that odd, special case, we'd have to remember to change the code back to the original version, which assumed a comma as the separator.

Our second revision makes a better way of doing this: When calling the function, we specify the string to split and the separating character.

• Second version: add a parameter for separator

```
FUNCTION rSplit
  *usage:    rSplit (name, ',' )
  parameters cString,  cSeparator
  jnSplitPos = at (cSeparator, cString)
RETURN alltrim(substr (cString,
  nSplitPos + 1 ))
```

But this still isn't as smart as it should be. We know that most of the time, the comma is the separating character; our function should reflect that. So we'll make the comma the default choice, unless another separating character is supplied to the function.

We can handle this default-or-optional feature if we can tell the function to find out how many parameters were passed to it. That's what the parameters( ) function does: counts how many parameters were supplied to the current function. If rSplit( ) has two parameters supplied to it, then a separating character has been explicitly defined. If it has one parameter, that means the function has to explicitly fill in what we let the user imply: the separator is a comma.

• Third version: make separator DEFAULT to a comma

```
FUNCTION rSplit
 usage:   rSplit (name)
 rSplit (offenses, ";" )
   parameters  cString,  cSeparator
   if parameters() = 1
       cSeparator = ','
   endIF
```

Continued on page twelve

On the Internet

Must-have sites

By Nora Paul
The Poynter Institute

Summertime and the living is easy. The Internet can help make some of your reporting tasks easy, too.

Since this issue of *Uplink*, according to our esteemed editor, is a hodgepodge, I thought I would honor the theme with a hodgepodge of Internet sites that you really should know about.

Locating statistics

Finding the right statistics to illustrate your story or to verify what a source might tell you can be hard. Here are two great sources to make locating statistics easy:

• FedStat: http://www.fedstats.gov/

"More than 70 agencies in the United States Federal Government produce statistics of interest to the public."

The Federal Interagency Council on Statistical Policy maintains this site to provide easy access to the full range of statistics and information produced by these agencies for public use."

In a beautifully organized site, this compilation of statistic pages gives you tables and charts on everything from breastfeeding to union membership from such sources as the CDC, the Bureau of Labor Statistics, the Social Security Administration. The search function lets you target the topic you need or browse through the alphabetical listing.

• http://www.sau.edu/cwis/internet/wild/Refdesk/Stats/staindex.htm#useful

"Where the Wild Things Are: Librarian's Guide to the Best Information on the Net How Much, How Many? Statistical Sources and Calculation Tools on the Net."

This directory, compiled by the St. Ambrose University library, links you to great statistic sources, evaluated for their usefulness and currency with the librarian's skills. Organized by economic statistics, social and demographic statistics, useful data and formulas.

Locating news stories

Getting ready to cover a story that has been covered in other papers?

Check out what the local coverage has been through the Net. Here are a couple of great sites that can make locating the local media easier:

• http://www.mediainfo.com/ephome/npaper/nphtm/online.htm

Editor & Publisher's Online Newspapers Listing — sure, there are several listings of newspaper Web sites but Steve Outing's continues to be one of the best, most comprehensive and well maintained. Find a local site and, hopefully, they will have a couple days of news stories available.

• http://sunsite.unc.edu/slanews/internet/archives.html

Newspaper Archives — I call them angels of the Internet, those people who compile useful listings of Internet resources and eliminate so much of the hunting and pecking.

One of the angels is Margot Williams, *Washington Post* researcher, who had her class at George Washington University find all the newspapers on the Internet that have archives of past stories available.

The handy chart lists by state the newspapers with archives, the date range available and the pricing (an amazing number still are free, get to them quickly, though!)

While you are at the Special Libraries Association, News Division site, check out another angel's listing — Kitty Bennett's, researcher at the *St. Petersburg Times*, compilation of expert directories and sources on the Internet: http://sunsite.unc.edu/slanews/internet/experts.html

Doing calculations

Quick, how do you convert meters to feet? Ha! I don't know either.

Check out this site to make hard calculations easier:

• http://www-sci.lib.uci.edu/HSG/RefCalculators.html

Calculators On-Line Center — "Links to over 3,750 calculators created by over 1,090 very 'CREATIVE' Individuals, Businesses & Tax Supported Entities World Wide." Find the calculator for meters to feet in Part II, Mathematics — Weights and Measures. The site it links you to will even convert to "football field" equivalents. Calculators On-Line Center is one of those sites that is just interesting to read through the possibilities available.

Directories & dictionaries

Having a virtual reference desk of basic books with facts, figures, dates and definitions

Continued on page twelve

Get more tips from Nora Paul by going to the Poynter Web site at poynter.org The site includes a list of upcoming seminars including "Reporting with the Internet" and also includes some CAR-related handouts.

Continued from page ten:

Language of your own

NICAR and IRE have several books that can help beginning CAR users and the old pros. Books include *The New Precision Journalism* by Philip Meyer and *100 Computer-assisted Stories*. For information on other books and ordering and pricing information, call Wendy Charron at (573) 882-0684 or send her e-mail at wendy@nicar.org

```
   nSplitPos = at ( cSeparator,
cString)
   RETURN alltrim(substr (  cString,;
nSplitPos + 1 ))
* end of rSplit()
```

The function seems perfect now: It's as smart and simple as it can be. Still, we're going to make one more improvement.

It may not even seem like a real improvement or a necessary one. It probably seems a bit arcane — an attempt to guard against the improbable.

I'm assuming three things about programs:

1) Nothing is improbable.

2) If the improbable happens, make sure that either there is no damage or there is a graceful way out.

3) The last improvement involves a significant programming concept.

Inside this function we are creating a new variable, nSplitPos, by assigning a value to it. We want to be able to use this new function in any new or old FoxPro script we create and we have to ensure that there are no surprises.

We may think that it is not very likely that there is a variable with the name nSplitPos already alive and working somewhere else in the currently running FoxPro script.

But if there were already such a variable, the assignment statement in our function would overwrite the current value held by nSplitPos.

Anytime your code does something you didn't intend or, worse, weren't even aware of, that's bad. While it may not be immoral, it can cause you real grief trying to find and fix the damage.

To make sure our function couldn't possibly interfere with any other variables, we'll add the statement:

```
private nSplitPos
```

This hides any other variables outside the function so it can't interfere with them.

• Fourth version: make sure variables created by the function do not clash with any others.

```
FUNCTION rSplit
   parameters cString, cSeparator
   private nSplitPos
   if parameters() = 1
      cSeparator = ','
   endIF
   nSplitPos = at (cSeparator, cString)
RETURN alltrim(substr (cString,;
   nSplitPos + 1 ))
* end of rSplit()
```

Richard Mullins can be reached at (573) 882-2127, or send e-mail to richard@nicar.org

---

Continued from page eleven:

On the Internet

is easy with the help of these great listings of ready reference sources on the Internet:

• http://lib-www.ucr.edu/infomine/reference/

Infomine: Scholarly Internet Resource Collections — click on one of the categories (Acronyms, Biographies, Calendars, etc.) and get a small selection of evaluated reference sources to help answer quick questions. Put together by the University of California, Riverside.

• ipl.si.umich.edu/ref/RR/

Internet Public Library — organized by type of reference work (almanac, dictionary).

• k12.oit.umass.edu/rref.html

Ready Ref. On the 'Net — arranged alphabetical by subject.

Going traveling

It's easy to find information about the places you will be traveling to this summer with the following, one of the best compilations of city information:

• http://local.yahoo.com/local

Yahoo: Get Local – this beautifully designed package of local information, links and news is browsed by putting in the ZIP code of the place you are going.

What you'll get back is information about the local weather, major league sports results, the searchable Yellow Pages, links to local news sites, and the regular type of Yahoo categories like: business, city guides, education, entertainment and arts, government, health, lodging, etc.

There is even a list of local UPI stories that can help you pitch your trip as a reporting excursion.

Have an easy summer, everyone.

Nora Paul can be reached at (813) 821-9494, or send e-mail to npaul@poynter.org

A Question of Diversity

By Jennifer LaFleur
San Jose Mercury News

After the 1990 Census, the *San Jose Mercury News* had charted the changes and breakdowns of our diverse area. A few years later, Miranda Ewell wanted to look at what has happened since then in the workplace and in our neighborhoods.

After testing a few different measurement tools, I decided to try applying a diversity index to our populations.

I chose the *USA Today* Index of Ethnic Diversity developed by Philip Meyer of the University of North Carolina at Chapel Hill and Shawn McIntosh of the *Dallas Morning News*. The *USA Today* diversity index eliminates the need to analyze the distribution of several factors with one probability-based index.

"The index has a range from 0 to 1, and its value represents the probability that two people chosen at random from the study population will differ along at least one ethnic dimension." (International Journal of Public Opinion Research, Spring 1992).

We chose this index because it seemed to both show trends of diversity and be easy enough to explain to readers in a few paragraphs.

Here's the formula to calculate the Index of Ethnic Diversity which we applied to each Census tract:

Step 1: Probability that two persons chosen from a population at random will be members of the same racial group: $P_R = (A^2+B^2+C^2+D^2)$ where A, B, C & D are the proportions in the population of whites, blacks, native Americans, and Asians or Pacific islanders.

Step 2: Probability that two persons chosen from a population at random will either be both Hispanic or both no Hispanic. This is a separate value because Census questionnaires ask Hispanic origin in a different question from race. Therefore it is possible to be both white and Hispanic or black and Hispanic.

$P_H=(H^2+N^2)$

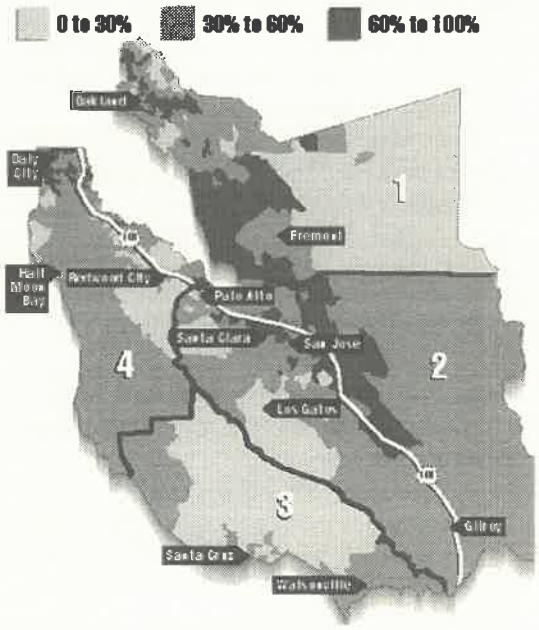
Step 3: Calculate the probability that two randomly chosen persons are the same in both race and Hispanic/non-Hispanic status: PR * PH

Subtract the result from 1: $1 - (P_R * P_H)$

We created a spreadsheet with the populations and put the formulas into the spreadsheet.

We applied the index to our local county totals, but also took our results by Census tract and mapped them to show the clusters of diverse areas.

Here's the result:



We also applied another measure to the counties of a whole over time: the Dissimilarity Index. This is another index from 0 to 1 — the closer to 1 the higher level of segregation in an area. For both Asians and Hispanics, the index has increased. The data also showed decreases for blacks; however, our county has too low a black population for most statistical measures to be accurate.

So why would you want to do any of this instead of general ethnic breakdowns for your area or even mapped for small areas like Census tracts. Those measures don't really show clustering over the mix of different populations. This would show low diversity for areas that are all Asian in our area or mostly white. This index provided a different look at our area. In addition this Index can be applied to other demographic measures, not just race and ethnicity.

Technology also changed the way we brought this story to the reader, in addition to showing maps in the newspaper, readers could go on-line, click on their neighborhood and see what the demographic breakdown is. This is something we just couldn't do on paper.

Jennifer LaFleur can be reached at
(408) 920-5000 , or send e-mail to
jenster@sjmercury.com

To view the on-line version of the *Mercury News* piece on diversity, go to www.sjmercury.com/news/local/census/index.htm

To order handouts about everything from using FoxPro for quick stories to using CAR to cover city government, call the IRE Resource Center at (573) 882-3364. Handouts from national conferences including the most recent NICAR and IRE conferences are now available.

Seminars, bootcamps, conferences

Get your training

Check out the NICAR Web site at nicar.org to find out times and dates for upcoming on-the-road training and bootcamps.

On-the-road training

NICAR provides specialized training in your newsroom.

Learn to transfer data from government files into newsroom PCs. Build spreadsheets for insightful stories on the beat. Comprehend documents with database managers. Navigate the Internet and on-line databases.

Cost varies. For information, call Lisa Barnes at (573) 882-8969, or to register, call the numbers below.

July 9, 1997 — New York, N.Y.
On-the-road training for CBS News

July 10, 1997 — New York, N.Y.
On-the-road training for *Newsday*

July 11-13, 1997 — New York, N.Y.
Computer-Assisted Reporting Workshop at New York University

Sponsored by NICAR and New York University. There are 16 spots available for this three-day hands-on seminar. Register online or call NICAR offices at (573) 882-0684 for a flier and registration form.

July 14-15, 1997 — Washington, D.C.
On-the-road training for the Paul Miller Program

Sponsored by the Freedom Forum.

July 17-18, 1997 — Chicago, Ill.
Join NICAR for hands-on training at the National Association of Black Journalists' convention in Chicago. Limited to conference participants.

July 25-26, 1997 — Philadelphia, Pa.
Reporting with Education Statistics Workshop at Temple University

The first of NICAR's specialized workshops on beat reporting, hosted by Temple University and *The Philadelphia Inquirer*. The workshop includes a day of panels that pair reporters experienced in education statistics with experts in the fields, and a day of hands-on training using the techniques they've described. Register online or call NICAR offices at (573) 882-0684 for a flier and registration form.

August 13-17, 1997 — Boston, Mass.
Join NICAR for hands-on training at the Asian American Journalists Association in Boston. Limited to conference participants.

August 19-22, 1997 — Huntington, W.Va.
On-the-road training for Marshall University

September 11-13, 1997 — Washington, D.C.
Computer-Assisted Reporting Workshop
Sponsored by NICAR and the Medill School of Journalism

October 20-24, 1997 — Ft. Lauderdale, Fla.
On-the-road training for the *Ft. Lauderdale Sun-Sentinel*

February 1-3, 1998 — Minneapolis, Minn.
On-the-road training for the Minneapolis Press Foundation

Conferences

NICAR will offer training and seminars at the following professional conferences, including the IRE and NICAR national conferences. Costs vary. For information or to register, call Lisa Barnes at (573) 882-8969

March 5-8, 1998 — Indianapolis, Ind.
NICAR Conference

June 4-7, 1998 — New Orleans, La.
IRE National Conference

Bootcamps

Bootcamps are week-long, intensive training sessions offered at NICAR's headquarters at the Missouri School of Journalism in Columbia, Mo.

As with on-the-road training, you will learn to transfer data from government files into newsroom PCs, build spreadsheets for stories on the beat, comprehend documents with database managers, and navigate the 'Net and on-line databases — but you'll be drilled all day, every day for a full week. Tuition ranges from $500-$1,000 depending on circulation or market size.

For information, call Wendy Charron at (573) 882-0684.

August 10-15, 1997 — Columbia, Mo.
NICAR Regular Bootcamp

September 28 - October 3, 1997 — Columbia, Mo.
CAR Workshop for Minority Journalists
Sponsored by NICAR and the Multicultural Management Program

January 4-9, 1998 — Columbia, Mo.
NICAR Regular Bootcamp

Growing collection of federal databases

From the NICAR library

NICAR offers a number of federal government databases. Here is a list of our growing collection:

• Federal Aviation Administration's accidents and incidents, including major plane crashes since 1974.

• NASA's air safety reporting system, including anonymous complaints by pilots and air traffic controllers. Useful for finding near misses and problems at local airports, 1988-1996.

• Federal Election Commission campaign contributions by individuals and political action committees, 1990-1997.

• The Health Care Financing Administion's 1995 database of all Medicare-funded inpatient work in U.S. hospitals.

• Federal Railroad Administration data for accidents, casualties, and highway crossings, 1991-1996.

• Coast Guard boating accidents, 1969-1994.

• Federal Aviation Administration data, including airplane maintenance work documented in the service difficulty report, pilot licenses and grades, and aircraft registration, 1974-1997.

• Home Mortgage Disclosure Act records, for tracking who gets loans and who gets turned down, and finding redlining patterns, 1992-1995.

• Federal procurement data, 1992-1996, includes breakdowns by agency.

• Alcohol, Tobacco and Firearms gun dealer records. 1993-1996.

• National Bridge Inventory System data, includes inspection grades, 1994-1995.

• FBI Uniform Crime Reports, a detailed compilation of crime data that includes statistical breakdowns of individual murders. This includes 1995 data.

• Social Security death records, by name and social security number, going back to 1937.

• Occupational Safety and Health Administration violation data includes worker accidents and exposures to hazardous chemicals by companies, 1974-1996.

• U.S. Department of Transportation truck accident and census data. It includes accidents by company and road.

• U.S. Small Business Administration loan guarantees, 1989-1995. This includes the name of the business, address, amount covered by the SBA, and status, including whether the loan went bad.

• U.S. Small Business Administration disaster loan guarantees, 1989-1995. This includes individuals and businesses, the amount covered by the SBA, and the status, including whether the loan went bad.

• U.S. Small Business Administration's list of minority companies certified for SBA assistance in seeking federal contracts. It includes the name of the company, its address, the owner, type of business and phone number.

• The National Inventory of Dams, 1991-1995.

• U.S. Department of Transportation hazardous materials accidents database, a collection of roadway, rail, air and waterway accidents from 1971 to 1996.

• U.S. Department of Transportation fatal accident reporting system. It includes all roadway accidents from 1975 to 1995.

• U.S. Coast Guard directory of U.S. merchant vessels. It includes the name of the ship, the managing owner, home port and various descriptive information.

For up-to-date prices and more information, call (573) 882-0684, or send e-mail to info@nicar.org

Need technical advice? Can't find what you need on the Internet? The bound edition of 1996 Uplinks is now available for $20 plus shipping. To order, call Wendy Charron at (573) 882-0684.

Bits, Bytes and Barks

Membership Message

You can join IRE via our Web server, or you can download a copy of our membership form in Adobe Acrobat format from the Web site and submit it by fax to (573) 884-5544.

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

NICAR Net

Recent topics on the IRE and NICAR listserv have included campaign finance data, troubles with traffic accident data, ethics, medical investigation, medical investigations and Access tips.

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

In the body of the message, type:

subscribe NICAR-L<your name>

subscribe IRE-L<your name>

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

Thanks for the Memories

Phoenix turned out to be a great conference thanks largely to the tireless efforts of volunteers and speakers. NICAR and IRE would like to send special thanks to speakers and demo room lecturers.

The conference would have been impossible to put on without the help of reporters, editors, researchers and professors across the country.

Plans are already being made for next year's conferences.

NICAR will hold its national conference March 5-8 in Indianapolis, IN.

The IRE National Conference will be held in The Big Easy – New Orleans, LA – June 4-7.

For information or to register, call Lisa Barnes at (573) 882-8969. We're also looking for any suggestions for improvements for next year's conferences and welcome any volunteers or speakers who want to participate.

Submit your ideas to lisa@ire.org

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

NON-PROFIT ORG.
U.S. POSTAGE
PAID
PERMIT NO. 309
COLUMBIA, MO. 65211