

Uplink

September 1996

A newsletter for the National Institute for Computer-Assisted Reporting

Newsrooms, start your engines

Uplink update

Don't stall while you're getting your CAR started.

In this issue of Uplink, newsrooms and reporters share how they got CAR programs up and running. The approaches vary from the *Seattle Times*, where Tom Boyer oversees a staff of researchers, to the *Minneapolis Star Tribune*, where CAR director Bill Loving trains reporters and runs a CAR station in the newsroom, and to WDIV-TV in Detroit, where Mike Wendland began CAR at home on his own equipment.

Also, Bill Dedman shares The Associated Press' guidelines for responsible Internet use, and Penny Loeb talks about *U.S. News & World Report's* approach to computer-assisted reporting.

NICAR's Richard Mullins tells how to get a Top 10 out of your database, Dave Migoya of *The Detroit Free Press* reveals a hot new database, and Maureen Smith of the *Star Tribune* shares how she and a team of reporters and editors used a spreadsheet on deadline to analyze school test scores.

Inside

On the Internet

Page 5

Calendar

Page 14

Bits, Bytes and Barks

Page 16

Newsrooms share CAR models

We did it our way

By Angie Gaddy

NICAR staff

Who does your data? That's what a lot of CAR newbies want to know from those using CAR in the newsroom.

Whether run by a database editor, the library staff or a computer-savvy reporter, newspapers are structuring a CAR hierarchy. Four large dailies share their CAR models:

Train them

It's been almost seven years since CAR began at the *Minneapolis Star Tribune*. At first, CAR was used in fits and starts, but in the past two years, a new push led by CAR director Bill Loving has gathered momentum.

Loving's goal is to have more than a dozen reporters skilled and using CAR on a semi-regular basis by the end of 1996. (See CAR Trips, page eight).

Loving is accomplishing this goal by training and coaching the paper's newsroom staff of 350, half of whom are reporters. He has trained at least 100 reporters to use the Internet in his own workshops and plans to set up workshops for tackling database managers. He has sent six staffers to NICAR bootcamps.

Loving has a CAR center in the newsroom equipped with two computers, including a Pentium and a 486 PC, a nine-track tape reader, a zip drive and some external hard drives. He plans to get a scanner and CD-ROM writer. CAR reporters build queries using FoxPro, Paradox and Quatro Pro, but most stick to Access' query by example to build their

own searches.

While Loving waits for reporters to build skills, he is focusing on acquiring databases. Loving acquires databases on nine-track tapes, via FTP, CD-ROMs and diskettes. Most of the databases are stored on CD-ROM in the library. He is just starting to put databases on the newspapers network server.

"We've made strides in the last year,"

Continued on page two

Broadcasters vie for support

Picture this

By John Sullivan

NICAR staff

If you're a broadcast journalist looking to start a computer-assisted reporting program at your station but can't get support from the higher-ups, you're not alone.

"I get an e-mail every day from reporters looking for advice on how to start programs at their stations," says Mike Wendland, author of "Wired Journalist" and CAR reporter at WDIV-TV in Detroit.

Wendland's own experience may benefit those trying to convince news directors of the benefits of computers in the newsroom. Wendland started with equipment he purchased on his

Continued on page seven

Uplink

September 1996

Volume 8, Number 9

A newsletter for the
National Institute for
Computer-Assisted
Reporting

Editor

Brant Houston

brant@nicar.org

Managing Editor

Jo Craven

jo@nicar.org

Assistant Managing Editor

Andrew Lehren

andy@nicar.org

Senior Contributing Editors

Richard Mullins

richard@nicar.org

Rosemary Armao

journa@mucmail.missouri.edu

Staff

Wendy Charron

Stan Dorsey

Angie Gaddy

Seth Hemmelgarn

Daniel Kim

Justin Mayo

Neil Reisner

John Sullivan

Wallace Winfrey

Copy Editor

Marilyn Lehren

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 \$35 (U.S.) annually, \$50 outside the United States.

Postmaster: Please send address changes to NICAR.

Send e-mail to

nicar@mucmail.missouri.edu.

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 a grant 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:

Newsroom models

he says. "We have an enthusiastic backing from the managing editor and executive editor. That's an absolute necessity."

An in-house homepage

One way to incorporate CAR into the newsroom is by using an internal newsroom homepage, says *Boston Globe* computer-assisted reporter David Armstrong.

The page — maintained by Matt Carroll, a part-time database editor, and by the paper's information services department — is accessible only in the newsroom, from the reporter's homes and from bureaus. It is loaded on newsroom Pentium PCs, which are all equipped with CD-ROM drives. While the paper converts its newsroom Atex system to PCs, a process that is half-finished, reporters can access the Internet and the paper's databases at two computers at a public station in the newsroom. Eventually, all reporters will have a PC at their desks linked to the paper's internal homepage.

Reporters can search databases of Massachusetts Motor Vehicle registration and driver's licenses; campaign contributions; and state, city and university system employees. The site also links to other homepages and newsroom chat groups. To make the system easier to use, there are front ends for accessing both the Internet and databases. Front ends are user-friendly interfaces that allow reporters to fill in templates, rather than build their own queries to maneuver around databases or the 'Net.

Armstrong says there's no push for a nine-track tape reader in the newsroom because he can download databases more simply and with fewer data errors through FTP. Now that more data comes on diskettes or CD-ROMS, nine-track tape readers, which are clumsiest, are going by the wayside in data transfer.

Access is loaded on all PCs, but most *Globe* CAR reporters prefer FoxPro. Excel is the most popular software program for those doing CAR — approximately 12 reporters use spreadsheets for stories.

The paper has brought in NICAR twice for training and has an in-house trainer from information services to help reporters. Carroll works one day a week updating databases,

acquiring data, answering reporters' questions about data and working on queries with reporters.

"It's not an ideal situation, because on the other four days a week he's unavailable to reporters working with him. It's a handicap," Armstrong says.

Armstrong, who is lobbying for a full-time database editor, says what the paper needs is "almost a cheerleader" in the newsroom, who encourages and sells the idea to reporters.

Make searching easier

With his desk equipped with a Pentium 90 with 1.6 gigs and DELL behind him, Mike Himowitz is literally situated between the newsroom and the library. His job as electronic news editor at *The Baltimore Sun* bridges library research with newsroom reporting.

Since CAR is just starting at the paper, Himowitz says the paper does not have a CAR model. Its's evolving, and Himowitz is discovering what works and what doesn't at his paper.

What works is creating front ends, or templates linked to database queries, to the paper's databases. Although it would make his job easier if reporters learned more CAR skills, his front ends makes it easier for reporters to find information without running to him with every question.

"Reporters don't have to become mavens (at CAR)," he says. "Some disagree. Some want to make reporters computer-attack dogs, but you can't expect every reporter to do it. Most will learn the skills that apply to their particular beat."

Reporters only need to type in the name or address they are looking for in the Baltimore city property tax records database, or the school name or county in the Maryland school scores database. They can access the databases, all stored on a server linked to the PCs, at a computer station outside of the newsroom library, equipped with four Pentium 90s, all with CD-ROM drives and hooked to the Internet. One computer is hooked to a scanner and an Overland nine-track tape reader.

There are two ends to CAR, Himowitz says: One is using it as an investigative tool. The other is using it to gather research for reporters. "We will see more response when more stories come out," he says. "The big payoff hasn't come just yet."

Continued on page three

CAR in any newsroom

Making it work

By Penny Loeb

U.S. News & World Report

In the three-and-a-half years I've been at *U.S. News*, we've done about 70 computer-assisted stories. They range from major projects on special education, bad blood, redlining and air safety to which Senator pays the most to fly home.

We've taught about 20 people at some level. They have been through 10 weeks of classes, gone to IRE conferences, learned on projects. We have one of the best computer setups anywhere, with a great amount of software, Web access and e-mail for everyone.

But there's a lot more to do. Here's what I've learned:

- A mission statement is essential. Everyone from the top editor to the person who runs tapes for data processing must endorse CAR.
- You need to let everyone know the possi-

bilities of CAR. But you need to meet them at their level. Some people will go through the classes and do projects. Some will need someone to download and clean data.

- You need to organize and prioritize. If you have a library, librarians should order and catalog data sets that are used routinely, such as contracts. You need to maintain lists of all data.

- You need to make the data user friendly. Guides for use help. You can also write front-end search menus in Visual Basic.

- You need a backup person who knows just about everything the CAR expert knows.

- The hardest thing to teach is a data mindset: the ability to integrate the possibilities for data into every step of a story and the perseverance to survive the inevitable snags.

Penny Loeb can be reached at (202) 955-2640, or send e-mail to ploeb@usnews.com

Continued from page two: Newsroom models

Start with research

Overseeing five news researchers, Tom Boyer began as database editor at the *Seattle Times* six months ago. The researchers look for news services, acquire new databases, coach, train and do higher end CAR work. Boyer's job is to coach researchers and reporters in CAR techniques, "as well as be 'stopper' on technical issues that no one else can deal with."

Networked Pentiums are scattered across the Atex system newsroom, all loaded with Excel and running on a TCP/IP network so the research staff can build World Wide Web front ends. The "hard-core" CAR computer is in the library. The paper's information systems department runs nine-track tape downloads, and they hire a local computer consulting company to handle tougher tape downloads. Boyer says they are buying a high-end database server, most likely an MS SQL server, to handle large files.

Presently, reporters come to researchers for help. But Boyer adds, "We do not search the *Times* database for staffers generally. If they need help, we schedule them for training."

As the *Globe* and the *Sun* have done, Boyer plans to set up a front end so reporters can run

their own database queries. And, as at the *Globe*, Boyer has created an internal Web site with a file of 3,000 source names, *Times* text archives and front ends to public records databases.

"We have to handle the basics first and organize things properly before I can really start CAR work in earnest," Boyer says. After that, he plans to start training reporters and editors on query techniques.

"But ultimately, it's not the responsibility of the database editor to make good CAR stories happen," he says. "If we're going to do this right, CAR has to come organically out of the work of the beat reporter."

Bill Loving can be reached at (612) 673-4493, send e-mail to loving@startribune.com

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

David Armstrong can be reached at (617) 929-3192, or send by e-mail to armstrong@globe.com

Tom Boyer can be reached at (804) 446-2478, or send e-mail to tboynew@seattimes.com

A complete list of handout abstracts, many dealing with CAR, from IRE's June National Conference held in Providence is available on-line at <http://www.ire.org/provfin2.html>, or from the IRE front page at <http://www.ire.org>, where you'll find a Resource Center link. Each file includes material distributed and an audio tape reference number for every session.

A complete set of handouts includes material from 80 different panels — a total of 1,075 pages. Cost for a complete set is \$150 and includes shipping.

Packets from individual panels are also available at the cost of 10 cents per page plus shipping. The minimum order is \$5. Nonmembers must pay an additional \$10 access fee.

To order, send e-mail to Dawn Hobbs at dawn@ire.org, call (573) 882-3364, or fax (573) 882-5431.

Setting 'Net policy

IRE, NICAR and The

Richmond Times-

Dispatch will offer the

Computer-Assisted

Reporting Workshop

for Minority

Journalists in

Richmond, Va.,

Nov. 15-17.

The hands-on training

will include navigating

the Internet,

spreadsheets,

database managers,

negotiating for data

and more.

Only 10 of 40 slots

remain available.

To register, call

Lisa Barnes at (573)

882-8969 ,

or send e-mail to

lisa@nicar.org

By Bill Dedman

The Associated Press

After collecting samples of newsroom policies on Internet use, The Associated Press devised the following policy to guide staff use of electronic resources. Some of the material is unique to AP, but may be useful to other news organizations interested in creating similar guidelines.

AP Guidelines for Responsible Use of Electronic Services

1. Accounts on AP's electronic services are intended to aid the business and professional activities of AP staff. The accounts are for business use only, the same as AP portable computers and cellular phones. If you have private business to conduct, use a private account. Although you can access the AP accounts from home, this is not a license to connect for personal use. Remember that connecting to electronic services uses a limited resource that costs the AP money. The same Internet connections and dial-up ports are needed to file and distribute AP text, graphics, photos, video, and data.

2. Each account is assigned for use only by the AP staffer. Sharing accounts is not permitted. No generic or departmental accounts will be assigned. Change your password monthly, choosing a password that would be difficult to guess (not your birthday, not a word in any dictionary).

3. Conduct business on electronic services as if you are appearing at a public meeting representing the AP, or writing a letter on company letterhead. After all, every message sent with an AP account is stamped "ap.org." What you write, even in private e-mail but especially in posts to lists and Usenet newsgroups, could be forwarded to millions of people, and no doubt will be saved somewhere by somebody. Many mailing lists that are erroneously thought of as private are routinely archived on Usenet or the World Wide Web, which are public. Even World Wide Web servers collect the addresses of all Internet users visiting them. And any user of AP's Internet server can see generally what activity any other user is doing. In short, if you wouldn't want your on-line activity to be shown on CNN or in

Times Square, don't do it on the Internet or America Online.

4. AP has long-standing rules against news employees participating in political activities or taking sides on matters of public debate. These rules apply to electronic communication as well. Do not express opinions about products, companies or individuals. Non-news employees, who may be unaccustomed to these rules, should remember that Internet readers won't know whether a user from ap.org is a newsperson. Even what a non-news employee does can reflect on AP's news gathering.

5. To do their work, AP staff need to participate in electronic discussion groups on professional or technical topics. Posting to other groups of general interest should be limited to seeking information. For example, a reporter doing a story on prostate cancer may post to a medical group, or a group for older men. Or a technician may seek help on a software discussion group.

6. When participating in discussion groups, be sure the reader knows that you are not stating AP policy. Someone reading a message from you@ap.org won't know AP's organizational structure. If complaints or questions come to you because you are identified as an AP employee, refer them to the appropriate supervisor.

7. Act as if the laws on libel and privacy apply to electronic communications. Remember that the laws of other jurisdictions may be more restrictive than your own. Respect the privacy of individuals, who may not be aware that their comments in electronic forums could be distributed by journalists. Do not quote private individuals or public figures from on-line communications unless you verify the identity of the author and assure yourself that the author meant to speak publicly. Often, it's best to contact people on-line, then to conduct an interview by telephone or in person. If you have on-line discussions to gather information, make sure the other party knows you are a working journalist. Although some Web pages and browsers allow sending of what's called anonymous e-mail, send only mail with your name and AP affiliation attached.

8. Apply the strictest standards of accuracy to anything you find on electronic services. The

Continued on page six

For an in-depth look at problems with databases, check out "Computer-Assisted Reporting: A Practical Guide" by Brant Houston. It can be ordered from NICAR or Investigative Reporters & Editors for \$26 plus shipping. Call (573) 882-2042.

Continued from page five:

On the Internet

laws, and an excellent listing of sources on journalism ethics.

WWW Virtual Library: Journalism

<http://www.cais.com/makulow/vlj.html>

John Makulowich, Internet trainer and early compiler of resources for journalists created this compilation of journalism related links. Section 8. Other Resources and Services, lists a number of research web pages.

Creating research pages

Think about the following when you go to design and compile lists of links. The answers to these questions will help give the page focus and clarity.

Planning and design:

- WHAT is the purpose of the web site?
- WHO is the audience for the site, who will be seeing the site?
- WHAT types of resources do you want to make available?
- WHAT kind of description will be made about the sites?
- WHO will evaluate recommended sites?
- HOW can sites be recommended?
- HOW will the sites be evaluated?
- WHO will maintain links to sites? What is the process for checking and maintaining?
- HOW will the site be organized? (around beats, topic areas, needs?)

• HOW will the site be searchable, how will be it navigated?

• SHOULD "non-Internet" resources be included? If so, which ones and how can they be?

Research Web features:

• Annotation: Why this site is useful, what the site contains, specific news applications.

• Instruction: Information about how to use the link (hints about searching).

• Searchability: How can you get to the exact site you need?

• Controlled vocabulary?: Is there a set of terms being used by the newsroom; if so, use them again!

• Links to non-Internetted resources: Pointers to internal databases, records

Guidelines for evaluating sites:

• WHO put the site/page together? What is their agenda? Why was it put together?

• WHO was the site designed for?

• WHAT does the site/page contain?

• WHAT perspective/side does the site take?

• WHEN was the information last updated?

• WHERE did the information come from?

• HOW is the information updated?

• HOW can this site help the journalist; what need is filled by this site?

• HOW easy is it to use (or how frustrating?)

Nora Paul can be reached at (813) 821-0583, or send e-mail to npaul@poynter.org

Continued from page four:

Setting 'Net policy

Internet is not an authority; authorities may use it, but so do quacks. Make certain a communication is genuine before relying on it as a source for a news story. More than one person may share an e-mail address, and e-mail addresses and Web page sponsorship can easily be faked. Ask yourself, "Could this be a hoax?"

Do not publish on the wire any electronic address without testing to see that it's a working address, and satisfying yourself that it is genuine. Apply, in other words, your usual news judgment.

9. Respect the copyrights of individuals and organizations, including the AP. Do not forward or post anyone's material without permission. Do not post or send to individuals

any proprietary AP material, including news stories, photos, graphics, audio, video, data, or any internal communication.

10. Abide by the courtesies of the electronic community. Courtesy requires basic technical competence. For example, be careful not to send a message to a mailing list that was intended for only one user. Don't type in all caps; people will think you're shouting. Avoid the "flame wars" that easily erupt when conversations are conducted on-line. And, because AP's Internet server has limited capacity, clean out your mailbox and home directory routinely, and log off when you're not using the system.

Bill Dedman can be reached at (212) 621-6984, or send e-mail to bdedman@ap.org

Continued from page one:

Support for broadcasters

own. "If you want to get something going, you just have to make the investment and buy the hardware to start at home. If you wait for someone to allocate the money, you may wait a long time," Wendland says.

Back in 1985 when Wendland started using computers to do stories, it was much harder to convince management to invest in hardware. In recent years, some producers have recognized the benefits of computers and made the top-down push to move number-savvy journalists into the high-tech fray.

Alan Cox at WCCO-TV in Minnesota admits his is an anomaly among broadcast programs. Cox says then news director John Lansing went to a 1993 IRE conference in Raleigh, N.C., and came back convinced computers were going to play a big role in the medium. Cox had done a county-by-county breakdown of state lottery funds in 1990 with Lotus 1-2-3 for DOS and a one-hour Macintosh training course at Kinko's. Lansing sent Cox for training, bought him some hardware, and gave him the luxury of two months free time to play around and learn what he could.

But Cox notes there was a precedent at WCCO. The station had done some of the earlier CAR stories under the direction of Pat Wyland. Cox produced stories right out of the gate and "Computer Center 4" was born.

Cox and Wendland offer some specific advice for starting programs:

If management is "strongly encouraging" you to embrace CAR, which may mean learn it or leave, Cox recommends getting professional training and as much material as possible. Cox says he came home from conferences with a briefcase full of IRE/NICAR handouts.

Both Cox and Wendland agree that if management does not push you onto the dance floor, you need to start moving and shaking on your own. Cox calls it the guerrilla approach: "Do whatever you can to get your hands on the equipment, even if it means going over to accounting and using what they have."

Wendland says a mistake many new CAR reporters make is biting off more than they can chew. He suggests starting small and producing stories to show news directors that investing in computers is in the station's best interest. Cox started out with Internet stories that showed

consumers the kinds of services that were available on the 'Net. The hard data stories came later. Wendland and Cox now do complex investigative stories using a variety of databases and software.

Hardware

Despite what many may think, CAR hardware necessary to do most stories can be bought on a reasonable budget. With memory as cheap as it is right now and ultra-fast processors making 486's hard to sell — and therefore a potential CAR bargain — a system is definitely within reach. In fact, Cox still uses a 486DX2-66 with 32 mb RAM and two hard drives of a gigabyte each. He says the memory is especially helpful for situations like election night, when he runs a graphics program along with mapping and spreadsheets.

Wendland's shop uses four computers and several laptops for their four-person CAR team. The computers are all Micron units ranging from a Pentium 200 to several 120s. The 200 has two gigs of memory, and Wendland uses a one-gig Iomega Jazz drive and a zip drive for larger jobs. They are all networked to a T-1.

Some hardware purchases are more specialized, such as a nine-track tape machine. However, new operations can get by without it since organizations like NICAR can process nine-track data.

Cox does have his own nine-track machine and some other useful tools, including an SVGA-to-NTSC converter that allows computer video to be put directly on air. Cox also says that after he lost a hard drive and found out how difficult it is to restore backups from nine-track, he got a Travan-style tape unit, which also comes in handy for moving large files from machine to machine.

Cox says his station uses an old 486 outfitted with an extra hard drive on the assignment desk so managers can search the station's databases.

Wendland has seen his CAR program grow to include a reporter, producer, and an associate-producer. Cox says he hopes his program expands with the purchase of some new hardware. Both programs prove that whether your resources are limited or unlimited or whether the initiative comes from you or someone else, you can do quality CAR stories for the broadcast medium.

Mike Wendland can be reached at (313) 222-0532, or send e-mail to mikew@wdiv.com. Alan Cox can be reached at (612) 330-2530, or send e-mail to wccotv@mr.net.

Mike Wendland is the author of "Wired Journalist: Newsroom Guide to the Internet." The book, intended for radio and television broadcasters, details many of the research tools available on the Internet, including an appendix that lists several hundred helpful Web sites. The book is available from:

Radio and Television News Directors Foundation.
1000 Connecticut Avenue, N.W.
Suite 615
Washington, D.C. 20036.
Phone: (202) 659-6510;
or e-mail: rtnd@rtndf.org

Count me in

By Maureen Smith

The (Minneapolis) Star Tribune

The next time you think about crunching numbers with a calculator, think again.

The education team at the *Minneapolis Star Tribune* walked into the land of computer-assisted reporting much like a nomad would wander into the forest: equipped with survival skills and resourceful minds, but little technical know-how. And we were more comfortable with calculators than computers.

But we got over it. On deadline.

It this summer. Our five-reporter-and-one-editor-team was strategizing around noon about the news to come: a breaking test score story (don't yawn yet). All we knew was that later in the afternoon, we'd have the test scores of scads of Minnesota eighth graders on new state-wide tests in math and reading.

Comparing school districts

We already knew that fewer students passed the basic skills tests than hoped or expected. And we knew that the students would be the first group in the state that had to pass the tests to graduate by the year 2000. What grabbed our curiosity was how districts compared. This was the first opportunity in Minnesota to eyeball scores of different districts on the same tests.

The state Department of Children, Families and Learning (the former Department of Ed, renamed and overhauled by the Republican Gov. Arne Carlson) was releasing data on each school district in Minnesota. We knew nothing about what the data looked like. We hadn't even thought of asking for it on disk.

By 3 p.m., a reporter came back armed with stacks of blue sheets of paper — one for each of about 300 school districts. Did it have the information the way we wanted it? Of course not. We wanted to know what percentage of kids passed and failed for each district.

Each blue sheet had the name of the district, the number of students who took the math and reading tests. But instead of telling us what percentage passed and failed, each sheet broke the students into six score categories that didn't neatly add up.

A passing score was 70 percent. The score categories were broken into students who scored 0 to 24 percent; 25 to 49 percent; 50 to 69

percent; 70 to 74 percent, 75 to 79 percent; and 80 to 100 percent. The number of students who fell into each category was listed. The percentage of students who fell into each category were also listed. But the percentages were rounded up, and didn't add up to 100 percent. In other words, we couldn't simply add up the percentages of kids in the three upper score categories (70 percent and above).

The bottom line was that we had to do the math ourselves.

Calculator vs. spreadsheet

Then began the debate: Should we use calculators, which we all knew how to use, or those unfathomable computer spreadsheets? Could we do either on deadline? Was it enough to have one reporter who knew how to use Lotus, another who had passing familiarity with the term "spreadsheet," and an enthusiastic but not technocratic editor? Our education editor, Maureen McCarthy, bargained on the risk. Her face lit up at the mention of a chart that would compare all school districts.

Fortunately, Bill Loving, since named CAR director, happened to be within earshot. "Hey Bill, do you think we could do this on a spreadsheet?" Loving and I sat down and I learned how to use Quattro Pro while we set up the spreadsheet. Another reporter who knew what a spreadsheet was prepared four others (including a copy editor) for data entry.

It was going on 3:30 p.m. as I agonized over figuring out the formula for percentage and finding the formatting command for a percentage column. I paged through Quattro Pro guides and played with formulas.

Meanwhile, three teams of two people each entered a stack of raw data: One read it, the other typed it in the computer. Each group also proof-read their data entry list once. One column for the district name, another for the district number. Six columns for the score categories.

By about 5 p.m., we had three lists of school districts and merged them together into a common spreadsheet that summarized the percentage of students who passed and failed at each district. In a column on the final spreadsheet, we added the kids who passed, or scored 70 percent or above. In another column, we calculated the

Continued on page fifteen

IRE and NICAR will hold a regional/ student conference Nov. 9 and 10 at Syracuse University. The conference will have sessions on paper and on-line records; interviewing, FOI, writing, and investigations on a budget for both print and broadcast. Cost is \$10 for members, \$50 for non-members, which includes a year membership worth \$40. The student cost is \$35, which also includes a year's membership worth \$25. Student members can also attend for \$10. Panels will take place on Saturday. Hands-on computer training will be available on Sunday for an extra \$25.

For more information, e-mail Rose Ciotta at ciotta@buffnews.com or Rosemary Armao at armao@ire.org

Today's Top 10 list

By Richard Mullins

University of Missouri/NICAR

I remember the first few times I was asked a certain "Is there a way to do this?" question. Unfortunately, I said there wasn't a way.

Turns out I was wrong (I'm writing this for penance). I should have said, "Not that I know of yet." I've managed to reassure myself by seeing my error as confirmation of one of the presumptions that I recommend to every reporter using a computer to commit journalism: If you can even dimly conceive it, there is probably a way to do it.

Here was the question: Is there a way to write a query that just gives you the top 10 (or 20 or 101) rows?

I'll cover Access first, then FoxPro. Apologies to those I've left out.

Access users who came to the software tainted with previous knowledge of SQL, a lingua franca of data analysis, probably noticed the extra keyword Access used: DISTINCTROW. I'll explain that one, too, after the ranking business.

In Access, this query, which orders candidates according to their fundraising, doesn't limit the result to any rank:

```
SELECT DISTINCTROW CandName, Receipts
FROM CandRept
ORDER BY Receipts DESC
```

Adding one keyword and a number gives the same results, but with a specified cutoff in rank:

```
SELECT TOP 10 CandName, Receipts
FROM CandRept
ORDER BY Receipts DESC
```

There is no opposite keyword for TOP to get the bottom of the list. Instead, you do that by specifying whether the ranked columns should be ordered ascending (the default) or descending. Using the TOP keyword without a corresponding ORDER BY clause won't give a syntax error, just an arbitrary selection of rows.

You can also add the keyword PERCENT to make the query return only a certain portion of the ranked data.

```
SELECT TOP 10 PERCENT CandName,
Receipts
FROM CandRept
ORDER BY Receipts DESC
```

The only way to specify this in an Access query is in SQL view. But you can use the Design View to quickly compose your query, then switch to SQL view and replace DISTINCTROW with TOP 33 or whatever.

In FoxPro, this SQL will return the top 10 fund-raisers:

```
SELECT CandName, Receipts ;
FROM CandRept ;
WHERE RECNO() < 11 ;
ORDER BY Receipts DESC
```

The trick here is in the WHERE clause and the built-in function, RECNO(). For every table and query result, FoxPro can figure out the total record count and the record number of every row. This information is not explicitly stored in the .dbf file; it's calculated on the fly.

So if you have ordered the query results by a certain column, the WHERE clause limits the result to the rows of the unfiltered result (the part you'd get without the WHERE) to include only those rows whose calculated record number (that's what RECNO() figures out) is less than 11.

Changing the expression in the WHERE clause lets you limit the results to a certain percent of the full set of rows.

```
SELECT CandName, Receipts ;
FROM CandRept ;
WHERE RECNO() < 20 * (RECCOUNT() / 100) ;
ORDER BY Receipts DESC
```

The built-in function RECCOUNT() determines how many rows are in the complete query result. Dividing that by 100 yields one percent of the set. The part you specify is the number following the less-than sign.

The last mystery to settle for this month is that DISTINCTROW business in Access.

In most queries, you can leave DISTINCTROW out, and it does not alter the results. It is only applied in queries where: 1) you are joining two or more tables; and, 2) the joined tables have a one-to-many relationship; and, 3) you are selecting columns from some, but not all, tables in used in the query.

Look up the word "TOP (SQL)" or "DISTINCTROW" in the Access on-line help.

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

Don't forget that the **Advanced Bootcamp** offered by NICAR and Philip Meyer, author of "The New Precision Journalism" and journalism professor at the University of North Carolina-Chapel Hill, will be coming up May 4-9, 1997. The seminar, which is held in Chapel Hill, will concentrate on stats and maps. For registration information, send e-mail to Wendy Charron at wendy@nicar.org, or call NICAR at (573) 882-0684.

Fire data is hot

The heat is on

The National Fire Information Reporting System is available from the National Technical Information Service at (703) 487-4650. The most recent year available is 1994, and, according to NTIS, the cost is \$1,265. If enough journalists are interested in this data, but find the cost prohibitive, NICAR will consider acquiring the database. Send e-mail to jo@nicar.org

By David Migoya
Detroit Free Press

The Detroit phenomenon known as Devil's Night was only weeks away, and familiar news-room rumblings of City Hall's dubious fire count accuracy began anew.

Each year, the day before Halloween becomes the prime focus of pyromaniacs intent on destroying the scant homes that remain in Detroit's more impoverished neighborhoods.

Similarly, each year, the *Detroit Free Press* puts a team of reporters on the streets to assess the latest fire damage and patiently awaits the release of the fire department's official fire count. In the first year of Mayor Dennis Archer's current term, the number of fires was the highest in several years—the first mar on what had promised to be a Phoenix administration.

Smoking

Despite the perceived frank numeric disclosure by city officials, editors wondered aloud whether the fire statistics were accurate or whether the true damage of Devil's Night, an event that, sadly, draws tourists from as far as Japan, was couched in the least terminal terms in an effort to spare scrutiny and embarrassment.

As in most cities, fire officials were not entirely cooperative in allowing reporters to inspect the city's documentation and, in the perceived absence of computer tapes, reporters often were left wallowing through a mass of near-indecipherable paper.

Enter the National Fire Information Reporting System, a little-known database that nearly every state keeps in nearly identical form.

I first learned of the NFIRS tapes while working with the *Belleville (Ill.) News-Democrat* in 1992. While investigating a home-sales scam in East St. Louis that kept a strangle hold on the predominantly black population for nearly 40 years, the city's fire chief told me that fires seemed more prevalent at the houses owned and sold by the two families of land barons that were my focus. Interested, I learned that East St. Louis, like many other cities its size, records fire runs in a giant log (too time consuming to analyze, I decided) and fills out

standardized reports for each of those runs. Those reports, I learned, are sent on to the state fire marshal, whose staff enters the information on a computer.

Windfall

Curious, I called the marshal and asked about the tapes. To my surprise, no one, least of all a reporter, had asked to see the tapes—ever. He was all too happy to share the information, an attitude I have found to be universal in every state I've worked since.

For a mere \$20, I received five years worth of nine-track tapes for the southwestern Illinois region. Every fire company, no matter if paid or volunteer, is required to fill out the reports that make up the database and send them to the marshal. The state primarily uses the information to assess fire fighting efficiency. I have since created a slew of other uses, much to the chagrin and delight of fire department officials.

In East St. Louis, I was able to determine there was a higher propensity for fires and a greater ratio for burning houses owned by the land barons than in the rest of East St. Louis (See related story, page eleven). Additionally, the database helped us find a flaw in state laws that allowed the land barons to collect the insurance proceeds, toss any buyers from the property without a dime of restitution, and leave the county holding the delinquent tax bills and a charred property for years and years. We found hundreds of such examples, all germinated from the database.

For the Devil's Night analysis, I obtained five years of data for the Detroit metropolitan regions (Tip: Get as much as you can for at least a five-year period) from the Michigan fire marshal. As in Illinois and Missouri, state officials were surprised someone was asking for the information. No one had ever done it before.

Unlike Illinois, the computer person responsible for keeping and analyzing the data was all too happy to provide it on diskettes (I have no nine-track reader at my City Hall office) and in a .dbf format (I prefer FoxPro). To my utter delight, it was free (This, unfortunately, has changed since word leaked out this summer that the database exists).

Continued on page twelve

Don't get burned

By David Migoya
Detroit Free Press

So, what's in National Fire Information Reporting System, this amazing, previously unknown database?

Simply put, everything.

Summarily: time, date, address, time call received, time dispatched, time arrived, time left the scene, number of rigs, which station, which apparatus, type of fire, locations within the building, cause, number of injuries and deaths to civilians and/or fire personnel, even the amount of water and length of hose that was used. House fires, car fires, grass fires, trash fires, false alarms, investigative calls. If there were a line for a firefighter running in place, it would be in there.

CAR dream

It's a CAR reporter's dream. It can also be his or her worst nightmare.

The data field code book is almost two inches thick and varies slightly from state to state. Some fields cancel each other out and others are completely unreliable. What at first glance appears to be a conflagration can quickly turn out to be a flaming paper bag. It has taken me four years to master the database because of its complexities. Best said, there's too much information, and in the hands of a haphazard reporter, it can easily turn into an ombudsman's career. It is extremely easy to make a mistake.

As with most databases, the information is entered by hand. Since 99 percent of the entries are numeric, keypunch accuracy quickly goes out the window. While the generic numbers — the number of house fires, the number of false alarms, etc. — are normally reliable, the more specific information should always be double checked on the database and with department records if you can get them.

Specifics

The information is specific. A house fire can be further detailed to show if it's a two-story, brick, no-foundation house. Buildings are detailed by their use, such as a bowling alley, a 22-unit apartment, a nursing home, an elementary school, or a church.

Addresses should always be confirmed through the department or in a city directory.

Firefighters often are reliable when it comes to dousing flames, but filling out reports is another matter. For example, the address of a burning house at the corner of an intersection will either be entered accurately or as "0000ELM" or as "ELMandMAIN," entries that make mapping difficult at best.

Car fires have no address so the entry may either be the intersection, the street block number or the address of the nearest house. Assuming the latter, the information at first appears to be a house fire. The car fire information is found in a different data field, and its type — whether it was an overheated engine, a smoldering seat cover or a true car fire — in yet another.

Smoke detectors

Last year, we analyzed the effectiveness of the Detroit Fire Department's two-year-old smoke detector giveaway program. The department didn't have the addresses to which they provided the detectors, but did have a tally by Census tract. Since the database also includes Census tract numbers, the story was still a go.

By isolating the number of actual house fires (you have to combine three data fields to assess the validity of a fire), separating the fires at abandoned houses, and adding the data fields that provide smoke detector information (there are three fields showing whether a detector was present, whether it worked and what role it had in warning occupants about a fire), we were able to tell which areas of the city had a greater number of detectorless fires and, arguably, the greatest need.

The database includes damage assessments, so we further isolated those fires that had significant damage. If there was little damage but a civilian injury, the information was included. Houses where there was more than one fire in a year were counted as one each year.

Other data faults

Then we grouped the information by Census tract, and to our dismay found scores of Census tracts that did not exist. Why? The department had been entering tract data based on indexes from 1980, not 1990 when several tracts were melded or separated into different tracts. We were forced to use a current city directory to fix

Continued on page thirteen

**IRE committees
always are on the
lookout for good
volunteers. Of
particular interest to
Uplink readers may
be the Computer
Committee, chaired
by IRE Director
Shawn McIntosh of
The Dallas Morning
News. To volunteer,
send e-mail to
McIntosh at
dmnproj@airmail.net.
For information on
other IRE
committees, send
e-mail to IRE
Executive Director
Rosemary Armao at
armao@ire.org**

Continued from page ten:

Fire data is hot

"100 Computer-Assisted Stories" is published by IRE and NICAR. Newspaper and television reporters recount how they got the story. The price is \$20.

To order, call (573) 882-2042.

Juggling numbers

After asking the Detroit fire department's computer analyst (they had computer info all along, but preferred a court battle to giving tapes to reporters) how the numbers actually are counted (they didn't know I had their database from the capital), I analyzed the Devil's Night fires for a five-year period. What I found was interesting. The numbers were indeed misstated, at least in non-fire department terms. (Tip: make sure you can at least get the same numbers a department reports; this way you know how they did it, and you can find their flaws.)

What wasn't reported were "exposure fires," or fires that occurred when a blaze from an adjoining building or house spread to another building/house/car, etc. To fire officials, this is one fire and is reported as one fire even though, in some cases, as many as five houses burn. If a garage burns and a car burns, too, then it is listed on reports only as a garage fire — even if the car is parked outside next to the garage.

As a result, we learned scores more houses burned on Devil's Night, upping the total from previous years. There was no slowdown at all. The number of fires remained constant and more homes were destroyed than the public realized.

Hot spots

Using Atlas GIS mapping software, we found the Devil's Night "hot spots." With the administration's latest effort to demolish 2,000 abandoned houses — the most likely Devil's Night target — before Oct. 20, we were able to visit those hot spots to see whether the program was working. We learned through interviews and observation that some of the most volatile areas were overlooked, primarily be-

cause the administration had not considered using its own database in the same manner.

Things changed quickly, and Devil's Night 1995 was one of the quietest ever. When numbers were released two days later, we were able to assess them accurately and fairly and provide the plaudits the administration deserved.

Who's in it

The NFIRS system is becoming more universal, although some departments still fail to file the reports, or provide them in only partially completed form. Unfortunately, there remain states that don't have the database at all, such as Oklahoma, where a medical study of the effectiveness of a smoke detector giveaway program could have been simplified had the database existed. We shared a similar story we did in Detroit and Oklahoma officials are pondering the best method to computerize their data.

Some states, such as Missouri, will personalize the data, adding information to the basic NFIRS and changing the acronym to reflect the change. There it is called MFIRS; in Illinois, it's IFIRS.

Although the database is available en masse from NTIS, be forewarned: It is expensive and cumbersome. For the tri-county area of Detroit, a single year contained nearly one-half million records. Remember: Every fire run (including the "save-the-cat-from-the-tree" variety) is reflected in the database. And with the number of fields extending from as few as 35 (the standard) to as many as 90 in customized systems, the workload increases exponentially.

David Migoya can be reached at (313) 222-6677, or send e-mail to migoya@det-freepress.com, or to 75613.262@compuserve.com

IRE-L and NICAR-L addresses

IRE-L and NICAR-L are located on lists.missouri.edu, a dedicated Unix system designed to service discussion groups run out of the University of Missouri. All posts to the

lists should be sent to: ire-l@lists.missouri.edu or to nicar-l@lists.missouri.edu

All commands for subscription changes should be sent to: listproc@lists.missouri.edu

Continued from page eleven:

What to look for

the tract data and used a computerized tract index provided by the U.S. Census Bureau to clean the information.

Since the department said it based its give-away on need, we used Census data to profile the tracts that received the detectors and those that had the greatest number of detectorless fires and greatest need. What we found was telling: The department often played politics with a \$15 smoke detector by providing them to areas that not only needed them least, but could afford them most. The areas that had the greatest need — and often the most injuries — received nothing.

To personalize the story, we isolated the detectorless fires that had an injury and crossed them with another database that provided the names of the injured (called a "casualty report," you must ask for this separately from the NTIS or state fire marshal). We tracked the victims, and told their story.

The department has since changed how it gives away smoke detectors, many of which are donated by large corporations. The National Fire Protection Association also has adopted our method as an effective way to assess smoke detector need, and recommends it to participating departments.

In East St. Louis, we found how unreliable address ownership data can be. There are two data fields — owner and occupant — that are generally blank, wrong or interchanged. We needed to isolate the fires at houses owned by the two land baron families. We did this by hand checking every address with the county's tax

rolls (the county had provided us with useless data tapes on three separate occasions, and we opted to do it by hand rather than waste any more valuable time). Once done, we were able to assess that the ratio of fires was greater for the land barons than in the rest of the city, indicating something more sinister might have been occurring.

Try it out

As you can see, the database is swamped with information. Some innocuous, some vital. Most is accurate; much is not. The only way to know is to get the database and play with it. Start with a simple story, such as the number of arsons (two codes fit the bill: suspicious and deliberately set) or the number of false alarms. Danger zone: These are initial reports. Get the final cause for every apparent arson from the local fire marshal; otherwise the data will be skewed.

Trying to tackle a much more difficult and complex piece, such as whether slumlords are burning their houses or whether the fire department effectively uses its resources, will do nothing but intimidate and frustrate you and cause you to fear what could be one of the newsroom's most valued resources.

If you run into a problem (which is likely), call the data analyst for the state fire marshal. They are the masters and often will be only too happy to help.

David Migoya can be reached at (313) 222-6677, or send e-mail to migoya@det-freepress.com, or to 75613,262 (CompuServe).

NICAR on track for NashCAR conference

NashCAR, NICAR's 1997 national conference, will be March 6-9 in Nashville, Tenn., with the support and help of The Tennessean and a panel of local journalists.

The conference will offer the widest range of panels and hands-on training yet. Building on the previous conferences in Santa Clara, Calif. (CAR Trek), and Cleveland, Ohio (CAR Rock), NashCAR will have sessions for beginners, intermediates and advanced users.

We'll have tracks for beat reporters, editors,

newsrooms managers, broadcasters, researchers and librarians, and trainers.

The conference will use of computer labs at local universities and will have a demonstration room to display equipment and software.

We'll also have vendors and numerous opportunities for journalists to find jobs.

In addition, we expect to explore the effect of the World Wide Web on journalism and how journalists are using the web for their stories.

For more information, call (573) 882-0684.

Get your training

Get all 12 1995 Uplinks in a bound edition for \$10 plus shipping. With the bound edition, you can review macros devised by the St. Louis Post-Dispatch's George Landau for cleaning data (January 1995); revisit advice from U.S. News and World Report's Penny Loeb on covering schools (April 1995) and tracking nonprofits (December 1995); and remind yourself of how other reporters have used OSHA, HMDA and more. To order, call Wendy Charron at (573) 882-0684.

On-the-road training

NICAR and, where indicated, The Associated Press provide 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.

Note: In the following list, "Open to all" means any journalist may sign up. "Closed" means the session is open only to members of the host organization.

- Richmond, Va., AP — Oct. 3-4, 1996. Open to all. (804) 643-6646.
- Kansas City, Mo., AP — Oct. 7-8, 1996. Open to all. All spots filled.
- Keene, N.H., Keene State College — Nov. 7-9, 1996. Open to all.
- Kiplinger Seminar, Columbus, Ohio — Nov. 8, 1996. Closed.
- Allentown, Pa., *Allentown Call* — Nov. 12-13, 1996. Closed.
- New Brunswick, N.J., AP — Nov. 15-16, 1996. Open to all. (609) 392-3622.
- Washington, D.C., *Washington Post* — Dec. 9-10, 1996. Closed.
- Washington, D.C. — Medill School of Journalism — Dec. 12-14, 1996. Open to all. (573) 882-0684
- Milwaukee, Wisc., *Milwaukee Journal* — Jan. 20-22, 1997. Closed.
- York, Pa., *York Daily Record* — Jan. 27-28, 1997. Closed.
- San Diego, Calif., San Diego Tribune and SPJ — Feb. 19-21, 1997. SPJ Day open to all. (619) 293-1261

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 Wendy Charron at (573) 882-0684.

- RTNDA (Radio and Television News Directors Association), Los Angeles — Oct. 9-13, 1996.

- SEJ (Society for Environmental Journalists), St. Louis, Mo. — Oct. 17-20, 1996.

- Multicultural Management Program-NICAR, Columbia, Mo. — Oct. 13-18, 1996.

- IRE/NICAR Student Regional Conference, Syracuse, N.Y. — Nov. 9-10, 1996. Cost is: IRE member \$10 for conference. Non IRE Member, professional rate \$50 for conference (includes \$40 IRE membership fee). Non IRE member, student rate \$35 for conference (includes \$25 student membership fee. Syracuse students, \$5 to attend conference. Additional \$25 fee for Sunday computer training.

- Minority Regional Conference, Richmond, Va., — Nov. 15-17, 1996. Open to minorities.

- Minority Regional Conference, Dallas, Texas — Dec. 7-10, 1996. Open to minorities.

- NashCAR, NICAR National Conference, Nashville, Tenn. — March 3-9, 1997. Costs are: IRE member, \$150 for entire conference. Non IRE member, \$190 for entire conference, including IRE membership. Renewal member, \$190 for entire conference, including IRE renewal. Students, \$125 for conference, including student membership. Late fee is \$15 for registrations not postmarked by Feb. 14, 1997

- IRE National Conference, Phoenix, Ariz., — June 12-15, 1997.

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

- Jan. 5-11, 1997. Waiting list only.
- Jan. 24-26 — Intermediate Bootcamp
- May 4-9, 1997 — Advanced Bootcamp concentrating on stats and maps at the University of North Carolina-Chapel Hill with Philip Meyer, author of "The New Precision Journalism"
- May 18-23, 1997.

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:

NEW • A monthly CD subscription for all 1995-96 Federal Election Commission campaign contributions by individuals and political action committees, plus all presidential matching fund requests.

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

- Federal Railroad Administration data for accidents, casualties, and highway crossings. 1991-1995.

- 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.

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

- Federal procurement data, 1992-1995, includes breakdowns by agency.

- Alcohol, Tobacco and Firearms gun dealer records. 1993, 1995.

NEW • 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 the new 1994 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.

- 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.

NEW • 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 1995.

- U.S. Department of Transportation fatal accident reporting system. It includes all roadway accidents from 1988 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 nicar@muccmail.missouri.edu.

NICAR's week-long

bootcamps in

Columbia, Mo., offer

hands-on training in

computer-assisted

reporting skills,

including the use of

spreadsheets and

database managers,

accessing data in

various media, such as

nine-track tapes, and

negotiating for data.

For more

information,

call NICAR

(573) 882-0684, or

send e-mail to nicar@muccmail.missouri.edu

muccmail.missouri.edu

Continued from page eight:

First ventures in CAR

percentage of kids who passed in each district.

Despite some technical obstacles between reporter's PCs and graphics Macintosh layout system, we found a way to send the whole spreadsheet through e-mail. Graphics artist Mark Boswell cleaned up the data and made a chart.

And we learned that obvious lesson in CAR-

land: Computers work and think faster, better, and more efficiently than calculators. They also come up with printed charts that are much neater than scribbling. And they're not as prone to human error. We plan to do more of the same in the future. It's easier than we thought.

Maureen M. Smith can be reached at (612)

698-6474, or send e-mail to

smithmm@mail.startribune.com

Bits, Bytes and Barks

Newsday gambling project

At *Newsday* in New York, we did a five-part series in December on how states had become hooked on gambling revenue and were introducing new games. We did a computer-assisted analysis of the demographics of who played the lottery, instant cash games and a new video keno game that you can play every five minutes.

To do this, we FOI'd the tape of more than a million lottery winners in New York State for the past five years and analyzed it using both FoxPro and a mapping program. We showed conclusively that the lottery was taking the most money from those least able to afford it. This contradicted the claims the state had been making to our reporters. Experts confirmed that this was a statistically valid approach, compared to just getting lottery revenue by outlet.

Other parts of the series included a profile of a major vendor, GTECH, a look at how states are spending heavily to promote gambling (\$40 million in New York), and a case study of Tunica, Miss., which has had problems since casinos went in. The reporters were Bill Falk, John Riley, Ford Fessenden and Stephanie Saul.

— Bob Tiernan, enterprise/database editor,
New York Newsday

Miami Herald on-line job

The *Miami Herald's* On-line Services Division is seeking to fill a key leadership position; the title and job description will depend on the skills and interests of the applicants, but the person hired will help manage this new department.

Among the experience that might be helpful: World Wide Web page design, computer-assisted research and reporting, graphic artistry, marketing and promotion, computer programming, newspaper copy editing and page design. The ability to write, edit and converse in both English and Spanish is a plus.

Send a resume and a note outlining why you think you're right for this job to Rich Gordon, the on-line services manager, at rgordon@herald.com, or via snailmail to him at 1 Herald Plaza, Miami, Fla. 33132.

NOTE: If you send via e-mail, do not send the resume as an attachment — embed it as text in your message.

ProfNet update

Dan Forbush, ProfNet sysop, says ProfNet queries can now target — or exclude — any of the following five types of institutions:

1. Colleges and universities: 700 institutions in 17 countries, including the entire Ivy League and Big Ten.
2. Corporations and PR agencies: 200 companies and agencies in all industry sectors.
3. Extended academe: 100 think tanks, national laboratories, and other government and non-profit organizations oriented to teaching, scholarship and research. Also includes PR agencies representing academically oriented organizations.
4. Government agencies: Includes 13 units of government at federal and state levels. (Elected officials not included.)
5. Non-profit organizations: 100 professional and trade associations, hospitals, museums, and libraries.

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

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