

Bits & Bytes



Data updates:

The IRE and NICAR Database Library has recently updated the following databases:

MAUDE

IRE and NICAR in early March updated its copy of the Manufacturer and User Facility Device Experience database, from the U.S. Food and Drug Administration, just as doctors in Baltimore and officials with another federal agency criticized a manufacturer for a faulty medical device. The MAUDE data includes information about problems that occurred with medical devices. Current as of 2001.

Home mortgages

This database includes a record of mortgage applications, including demographic information about loan applicants, including race, sex and income; the purpose of the loan; whether the buyer intends to live in the home; the type of loan; and the outcome of the loan application.

The dataset also includes geographical information on applicants, such as Census tract, metropolitan statistical area, state and county. Current as of 2000.

continued on page 2

TERRORISM

Building it from scratch: WTC victim database

By Paul Overberg and Anthony DeBarros, *USA Today*

Some of the best databases are not the ones you buy or acquire through FOIA but the ones you build yourself, one fact at a time. That's what *USA Today* found when it set out to chronicle the thousands of lives lost in the terrorist attacks of Sept. 11. The result was several stories that effec-

tively challenged official reports of the death toll, as well as a major enterprise story that chronicled the last moments inside the World Trade Center.

On Sept. 12, *USA Today's* four database editors huddled to catch their breath and draw up a CAR-based plan.

continued on page 18

SPOTLIGHT: WORKPLACE

Mining workplace data

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

A quick scan of your daily paper or network news program shows that workplace woes are big news: Enron, once a little-known energy company, now is a household word. Workers – and not just at Enron – are fretting about the health of their employer-backed retirement savings plans.

In this issue of Uplink, we focus on workplace data. Stephanie Armour of *USA Today* recounts how she used computer-assisted-reporting's most basic tool – the spreadsheet – to build her own database of cases of immigrant maids being abused by their employers. In another article, Armour shares some reporting ideas and data

SPOTLIGHT:

FOR MORE ON WORKPLACE SEE:



- Following the riverboat pilots, p. 2
- Immigrant employee abuse, p. 4
- Workplace data resources, p. 6
- Beware of OSHA data, p. 7

sources that you can use to scrutinize employers and suspect workplace practices.

Over the years, many journalists have plunged into the massive enforcement database kept by the federal Occupational Safety and Health Administration, only to come up gasping for air. Ron Nixon, the national trainer for IRE

continued on page 26

Bits & Bytes

continued from page 1

Trucks

The truck accidents database consists of accidents that involve commercial vehicles that weigh more than 10,000 pounds, as reported by state transportation agencies to the U.S. Department of Transportation. The vehicles included in this database are buses, semis, moving trucks, rental trucks and hazardous-material carriers. Current as of 2000.

Fatal accidents

The latest updated Fatal Accident Reporting System database consists of three relational tables, containing data on automobile accidents on public U.S. roads that resulted in the death of one or more people within 30 days of the accident.

In the past, the NICAR data files were available in four relational tables. For the current update, NICAR used SAS software to replicate the data tables used by the Department of Transportation and translated those tables into DBF format, which can be opened by almost any database manager software.

In the latest data update, SAS translates the official DOT codes to plain English in the data tables. Covers 1975-2000.

For more information about the database library's collection, including purchase prices, go to www.ire.org/datalibrary/databases/. To order data, contact NICAR at 573-884-7711 or download an order form at www.ire.org/datalibrary/orderform/orderform.pdf.

SPOTLIGHT: WORKPLACE

Following the riverboat pilots

By Jeffrey Meitrodt, *The New Orleans Times-Picayune*

Teal Grue was the poster boy for our lead story about what's wrong with Louisiana's archaic system of selecting and overseeing Mississippi River pilots.

Grue dropped out of high school after 10th grade. He pleaded guilty to drunk driving in 1990. He tested positive for marijuana in 1998 and was forced into drug rehab. And he didn't get his first maritime license until 1999, less than two years before his father helped him join the elite ranks of Louisiana's river pilots.

We didn't write about Teal Grue because someone tipped us off. His stunning resume was simply one of many we re-

Mississippi upset a lot of people. So did the rest of "River Barons," a four-day series by maritime reporter Keith Darce and myself that started running in *The Times-Picayune* on Nov. 4, 2001. State legislators have called on the governor to hold a special session to deal with the issue in 2002. The governor, who stayed out of previous reform efforts, has finally agreed that the system needs fixing.

We didn't think the issue was that provocative last spring, when we first decided to take a look at the people who guide foreign-flag vessels along the Mississippi River. At the time, we

	education	edu notes	college degree	grad date	arrested	convicted	crime notes
1	college grad		yes	5/1/1999	no	no	
2	some college	1 year college	no		no	no	
3	high school grad		no		no	no	
4	high school grad		no		no	no	
5	some college	2 yrs college	no		no	no	
6	college grad	1 year law school	yes	12/22/1995	no	no	
7	high school grad		no		no	no	
8	high school grad		no		no	no	
9	high school grad		no		no	no	
10	didn't grad HS	GED (finished 10th grade)	no		no	no	denied all crime data
11	some college	1.5 yrs college	no		yes	yes	admitted to two DWI convictions
12	college grad	master's degree	yes	8/20/1999	no	no	
13	high school grad		no		no	no	
14	didn't grad HS	GED (finished 10th grade)	no		no	no	
15	didn't grad HS	GED (didn't finish HS, got GED)	no		yes	no	charged with "concealed weapon" age
16	high school grad		no		no	no	
17	didn't grad HS	GED (finished 10th grade)	no		no	no	
18	didn't grad HS	GED (didn't finish HS, got GED)	no		yes	no	lied about arrest; outstanding warrant
19	didn't grad HS	GED (didn't finish HS, got GED)	no		yes	yes	lied about arrests

viewed and typed into an Excel spreadsheet after *The Times-Picayune* decided to investigate what some believe is the ultimate good-old-boys club in Louisiana.

River pilots have one of the most dangerous jobs in the state. They operate on the most treacherous waterway in North America. And under state law, they are free to elect and regulate themselves – a privilege shared by none of their peers in the United States.

The idea of Teal Grue steering a 100,000-ton oil tanker down the lower

wanted to find out if two rumors were true: that Louisiana pilots are some of the highest paid mariners in the country and that widespread nepotism makes it virtually impossible for non-relatives to join the ranks.

Getting started

The first task was relatively easy. Since their pay is typically decided by a public agency, pilot compensation is a matter of public record in many states. To fill in the blanks, we called around to various pilot groups and interviewed experts who had conducted compensation studies.

We found out that Louisiana pilots, who earned about \$321,000 each last year, were near the top of the heap.

The second task was much harder. Unlike most states, which regulate pilots through independent state agencies, Louisiana leaves that job to the pilots themselves. That means that most of the records we needed to document the prevalence of nepotism, and investigate other pilot practices, were in the hands of the pilots, who are legally considered state officials.

We knew that the quickest way to kill our project was to file a voluminous public-records request at the beginning of our reporting. In the past, the pilots have vigorously resisted outside scrutiny. One of the groups even refused to give a copy of its bylaws to the National Transportation Safety Board when the federal agency was investigating an accident involving one of the group's pilots.

To avoid litigation, we took a low-key approach. We contacted the leaders of the three pilot groups that operate on the Lower Mississippi and asked them to help us understand what they do and why they're paid so much money.

The groups were happy to cooperate. They had just taken a bruising in the press over the huge raises they had won from the Louisiana Public Service Commission, and they were also smarting from a failed legislative effort last spring to shift pilot oversight to the state.

The pilots let us ride with them on ships, bunk at their stationhouses and interview anybody we wanted in their organizations. Two months later, when we filed our first records request, the pilots were unhappy but still cooperative. By then they understood that we were as interested in what they do and how they do it as we were in nailing them for their mistakes.

The records

The main records we wanted from the pilots were their job applications and the accident reports they filled out if one of their ships clipped a bridge or smashed into a fleet of barges.

None of the pilot records were available electronically. We wound up with six boxes of documents, and it took us about three weeks to type everything we needed into Excel spreadsheets.

In the past, the pilots have vigorously resisted outside scrutiny.

The applications were rich with revealing personal information. All applicants were required to furnish a birth certificate, and many were also asked to supply the names of their siblings and spouses. That proved invaluable in documenting nepotism, as well as tracking campaign contributions (many of which were made by pilot wives). To help us summarize the data, we added yes/no fields indicating whether an applicant's father was a pilot, or whether any other family member was a pilot.

The results were irrefutable: Of the 100 people elected to become river pilots in recent years, 85 are related to other pilots, with some able to trace their family connections back five or six generations.

The applications also yielded surprising information about the criminal and educational backgrounds of many of those selected to become pilots. We created separate, standardized fields for each type of data. On education, for instance, we used four descriptive phrases: "didn't grad HS," "HS grad," "some college," "college grad." Anything interesting, like the fact that someone dropped out of school after 10th grade, we threw into an all-purpose "notes" field.

The analysis

We analyzed the data using both Excel and Access. If we wanted to know, for example, the number of people arrested for drunk driving, we'd export the spreadsheet containing application data into Access and run a query to count the number of applicants ar-

rested for DWI. If we just wanted to find an example of someone with a DWI, we would sort the data in Excel by a yes/no field and look at the "notes" column to see who provided the best anecdote.

Adding that kind of summary information to the spreadsheets saved us time, since we didn't have to go back and look at all of the hard records every time we wanted to find an example for a story.

We also added a field showing the number of years an applicant had been working in the maritime business. This wasn't a question on the form, but we could calculate the figure by looking at various records (resumes, character references, etc.) contained in the application file. This made it possible for us to prove that winning candidates were invariably less experienced than the outsiders who tried to break into the organizations.

Surprising results

One of the most unexpected findings of our serious involved pilot discipline. Originally, we didn't think public safety would be a major factor in the series. But to make sure we didn't miss anything important, we asked the groups to provide accident and pilot misconduct records, which are investigated and judged by the pilots themselves.

The files were a mess. Names were frequently abbreviated, and rarely spelled the same way twice. To solve dirty data problems, and make it possible for us to combine data from various spreadsheets, we assigned each pilot a unique identification number. That made it easy analyze the data — to calculate, for instance, which pilot had the most collisions on the river.

The records contained several bombshells. First, we discovered that two pilots were allowed to keep their jobs after testing positive for cocaine use in 2000. Second, we found that one group racked up nearly 100 accidents yet failed to discipline a single pilot in a decade — a fact that dismayed even the group's staunchest supporters in Louisiana.

Considering that there is an average

continued on page 4

COMING IN JUNE

think SMALL

Interested in news about
America's growth engine:
Small Business?

Please join SABEW for our
second Small Business Reporting
Workshop on June 13-15 at
Wyndham Miami Beach Resort.

Registration fee is \$50 and
includes housing, conference
materials, meals, field trips and
writing critique. Minority
scholarships available.

Hurry! Attendance is limited.
Details at www.sabew.org

SABEW

Society of American Business Editors and Writers, Inc.

Workshop underwritten by the Kauffman Center for
Entrepreneurial Leadership at
The Ewing Marion Kauffman Foundation

tracking

continued from page 9

person. An interesting side note was that all of the spouses we found listed "homemaker" in the occupation field.

By doing these painstaking queries I found things my initial match missed. For example, I had an executive listed as J. Clifford Baxter (the one who later killed himself). In the FEC data he's listed as John C. Baxter. The join query I did would not find this. In addition, some of the executives listed other corporations as their employers. Further research revealed these were the correct people and often the other company was a subsidiary of Enron or worked closely with them (e.g. Vinson & Elkins law firm).

Once I had flagged all of the contributions made by the Enron executives, I pulled this chunk out of the data and started the analysis. We asked all kinds of questions: How much went to Bush? Who were the other big recipients? How much in soft money? How much in hard money? (Full disclosure: We ultimately used soft money figures from the Center for Responsive Politics Web site.)

Because of time constraints, we chose not to hunt down all contributions from Enron employees. This would have required far more time because of the dirty data issues noted above. We did the entire analysis and reporting in five days — which included three solid days of data analysis and checking.

Fast results, response

The data work for this story was not difficult, just a little time-consuming. I used Visual FoxPro for most of the work (since that is what I'm most comfortable with), but then kept track of results in Excel. I then used the spreadsheet for figuring percentages and totals. It was also a good way for me to have results readily available for when reporters John Dunbar and

Robert Moore or editor Bill Allison needed to see what was going on.

The results of our story were swift and exciting. The day after we posted the story on our Web site, Attorney General John Ashcroft recused himself from the investigation after Sen. Henry Waxman showed him our story saying he had received \$55,000 from Enron.

MaryJo Sylwester, formerly the IRE and NICAR Data Library administrator, can be reached by e-mail at msylwester@publicintegrity.org

readme.txt

The Center for Public Integrity's Enron stories are available at www.public-i.org

The first story, which documented campaign contributions and stock sales by executives, is specifically located at www.public-i.org/story_01_010902.htm

Where to get FEC data

NICAR:
www.ire.org/datalibrary/fec
FEC:
www.fec.gov/finance/newftp1.htm
FECInfo:
www.tray.com (subscription required)

You can also search campaign finance data at

FEC:
<http://herndon1.sdrdc.com/fecimg/query.html>

Center for Responsive Politics:
www.opensecrets.org

The National Institute on Money in State Politics lets you search for state campaign finance data at: www.followthemoney.org

TERRORISM

How data can correct government stats

By Mark Fazlollah and Peter Nicholas, *Philadelphia Inquirer*

There was quick action when the *Philadelphia Inquirer* reported that the Department of Justice had dramatically overstated its record of convicting terrorists, with conservative and liberal congressmen calling for an investigation.

The December story, which documented a five-year pattern of inflated terrorism statistics, went unchallenged by the Justice Department. The U.S. attorney's office in San Francisco, which had one of the worst records of pumping up the stats, acknowledged almost immediately that its numbers were wrong.

Among San Francisco cases of supposed terrorism were: the Mexican immigrant who concocted a phony passport application; the babbling man who walked into an FBI office and threatened to kill former President Bill Clinton — though he didn't realize Clinton was no longer president; the woman who totally lost it on a flight from Australia and demanded that the stewardess bring her the Virgin Mary (not to be confused with a Bloody Mary)

It was terribly tedious, but it produced results.

There were many other run-of-the-mill criminal cases that federal prosecutors were listing as terrorism. And the misinformation was going to Congress and the public.

Misrepresenting figures

"Where they were misrepresenting those figures of the terrorists convicted, it shows they may be lying to the Congress of the United States,

and those responsible for that should be removed from office," Rep. Dan Burton, R-Ind., chairman of the Committee on Government Reform said in a news conference called in response to the *Inquirer* story.

Burton's committee pulled the General Accounting Office into the battle to get an accurate accounting of terrorism cases. The GAO says its investigation probably will be ready in this spring.

Here's how we got the story:

Our basic data came from Syracuse University's Transactional Records Access Clearinghouse.

Each year TRAC gets raw Justice Department data and markets it in a highly useable form. The Justice Department organizes all federal cases into different criminal categories ranging from bribery to brutality and from white-collar crime to witness intimidation. And there is terrorism.

The justice data was complete up to Sept. 30, but it did not include the names of defendants or their federal court case numbers.

Finding a sample

To get that, we paired the justice information with federal court data offered through the Public Access to Court Electronic Records (PACER). This is a commercial service but the charges are reasonable — often only 25 cents to electronically access a full court docket for an individual case. The pairing is painfully time consuming and best suited to small samplings of cases.

The justice data listed only 463 cases of terrorism. We further reduced the number of cases by focusing only on

readme.txt

About the data

The Transactional Records Access Clearinghouse (TRAC) is a data gathering, data research and data distribution organization associated with Syracuse University.

Established in 1989 as a research center at Syracuse, it also has offices in Washington, D.C.

TRAC provides comprehensive data about federal staffing, spending, and the enforcement activities of the federal government. Its service TRACFED allows journalists to explore the database of U.S. attorneys' cases, reported by the Justice Department.

For more about TRAC, go to trac.syr.edu

federal convictions that were labeled as terrorism cases.

The common denominators in the two data systems are the dates of filing criminal charges, the federal court districts in which the cases are filed and the lead criminal charges.

PACER allowed us to select all cases filed on a specific date in any federal court district. For each federal district, we entered the date of each Justice Department's terrorism case.

Digging into files

The northern district of California, for example, had 14 convictions over a five-year period, requiring us to review the electronic docket information for every case on each of the 14 days.

It was terribly tedious, but it produced results.

Usually, there were no more than a dozen criminal cases filed on a single day. Of those, there usually would be

continued on page 12

readme.txt

Stories

The stories mentioned in this edition of Uplink will soon be available from the IRE Resource Center. Search the story database at www.ire.org/resourcecenter for order numbers. Call 573-882-3364 to order these and other investigative stories.

Databases

To view the list of databases available at the IRE and NICAR Data Library, view the Web site www.ire.org/datalibrary. From this page, you can also download an order form. The library's telephone number is 573-884-7711.

Tech tips

Have you found a technical solution that saves you time? Share it as an *Uplink* Tech Tip, a story of 500-1,000 words showing a technique that you'd be willing to share. Please contact Managing Editor Jeff Porter at jeff@nicar.org or by telephone at 573-882-1982.

Training material

For a calendar of IRE and NICAR training events, visit the Web page www.ire.org/training/otr.html.

Tipsheets

The IRE Resource Center offers a searchable database of tipsheets for journalists at www.ire.org/resourcecenter. Many are downloadable, or call 573-882-3364 to order tipsheets.

data

continued from page 11

only one with the specific criminal charge matching the Justice Department data.

There were additional checks we used to ensure we identified the right cases. PACER gave us the names of the judge and federal prosecutor in each case. The Justice data provided the initials of the judge and prosecutor. Once we had the name of the defendants, we could look for news articles that had been written by local papers.

We contacted defense lawyers, who were shocked that the Justice Department was classifying their clients' cases as terrorism. They faxed us copies of indictments and other court records. Reporters from our sister Knight Ridder papers pulled court files for us in several federal districts.

Reactions, examples

Confronted with the results of our research, some federal prosecutors were willing to discuss cases. We also reached out to FBI agents that investigated them.

In one supposed case of terrorism in Seattle, two commercial pilots were feuding.

Pilot Eugene Allen tried to hurt the other's career by sending authorities a phony report that the pilot was planning a hijacking. Officials quickly called in the FBI's anti-terrorism unit in Seattle.

The lead FBI agent, who worked three years with the anti-terrorism task force, cracked the case in two weeks. He got a confession from Allen, who eventually pleaded guilty to misdemeanor charges and was sentenced to a brief period of probation.

Reality of terrorism

The agent told us this was his "only real case of terrorism" in three years. Even after Sept. 11, he didn't understand the reality of terrorism.

Perhaps that was the most disturb-

ing problem we uncovered.

Now Congress wants to know how pervasive the problem is.

Other ways the data can be used: The same test that we did with the Justice Department's terrorism reporting can be done with other categories of crimes.

For example, are federal environmental enforcement prosecutions hitting big polluters or only getting small fish? Was it a mechanic who accidentally dumped a barrel of oil or was it a petroleum company that polluted Houston bay?

The same can be done for analyzing enforcement of organized crime cases, anti-trust cases, medical fraud, money laundering or any other category of federal prosecution.

Even for large categories of cases, such as drug enforcement, reporters can do sampling.

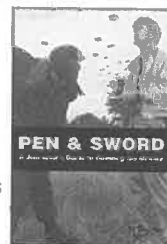
Mark Fazlollah can be reached by e-mail at mfazlollah@phillynews.com.

Peter Nicholas can be reached by e-mail at pnicholas@krwashington.com.

Pen & Sword

A Journalist's Guide to Covering the Military

Let veteran reporter Ed Offley be your guide to the military beat. Offley has covered military issues and conflicts for *The Seattle Post-Intelligencer*, *The Stars and Stripes* and other publications for more than 20 years.



Offley's clear, concise writing in *Pen & Sword* will help you understand:

- Military basics all reporters should know
- How to organize a military beat
- Protocol for interviewing military personnel
- How to go into the field with military units
- How to deal with the Pentagon
- How to prepare for war
- How to cover the defense industry

\$24.95. Softcover, 304 pages. ISBN 0-9665176-4-4

Order on-line: www.marionstreetpress.com
or call Ed Avis at 708-445-8330.

Marion Street Press, Inc.,
P.O. Box 2249, Oak Park, IL 60303

Q&A

From the Lists

Listserv messages you may have missed between members of NICAR-L, Census-L and IREplus-L. To learn about joining these mailing lists and searching their archives visit www.ire.org/membership/listserv.html.

Downloading with FTP

Q: What does using FTP software do for you in downloading that simply right-clicking and saving doesn't?

A: I had some problems downloading files through Netscape; when I started using the FTP client, not only did those problems disappear, but I also started seeing about 30 percent faster download times. It also allows you to select an entire set of files to download, rather than right-clicking on them one at a time. If it's just one file, I'd probably right-click on it, too. If it's a set like the 40 per-state SF-1 census files from last summer, right-clicking gets a little tiresome.

SPSS and Excel

Q: I don't use SPSS and know nothing about it. But I want to tell someone who has some data in SPSS how to give it to me in Excel (I'm not sure they know for sure how).

A: File/Save As/Excel ought to do it, as long as the SPSS table doesn't have more variables or cases (i.e., columns or rows) than Excel can handle.

Copying import specs

Q: I was under the impression that you could export or somehow copy import specifications between Access databases, but I can't remember how to do it nor find any way to do it in the help file. Can someone explain how to do it or point me to the relevant help entry?

A: Open the database to which you want to add the import specs. Go to File, Get External Data, Import. Then select the database that contains the import specs you want. On the import dialogue window, click the button labeled Options. Check the box labeled Import/Export Specs. Then click OK. It won't seem like anything happened, but the specs should be in the new database.

Summarizing with maps

Q: I need help creating a dot-density map. I have a point (dot) theme. Let's say each dot represents a single crime in the county. And I have a census tract theme. How do I assign the dots to their respective tracts and then tell ArcView to show me one dot for, say, every 50 in each tract?

A: If I understand, you already have a dot density map with one dot per crime, but want to aggregate it by tract and restate it as 1 dot per 50 crimes. ... If this is the case, you want to take you're joined table and click on the summarize button, which is a Sigma... this will start a dialog box, through which you want to select your TractId field and Summarize by Count. ... What you'll end up with is a .dbf that has:

```
tract 1 504
tract 2 444
```

Then you can add this new .dbf table and join it to your original tract theme and make a dot density map (double-click on the theme and pick dot density instead of graduated value).

Splitting those names

Q: I have a "name" field in an Access table I need to separate. The names appear like: SMITH, JOHN DAVID. I need to get the individual names in to different columns via update query... simply tell Access to put all characters before the comma in the first new field (lastname), and all remaining characters after the comma into a second new field (restname).

A: Use instr() to find the comma. Something like this:

Left(name,instr(",")-1) for the last name. Mid(name, instr(",")+1) for the rest.

continued on page 15

ONLINE

Handy Internet resources for reporters

John Tedesco, *San Antonio Express-News*

What kinds of Internet sites are being bookmarked these days by web-savvy reporters?

I asked members of NICAR's listserv recently for the most helpful sites they've come across and quickly learned that journalists have staked out interesting and informative territory on the World Wide Web.

The listserv itself, and the replies I received, became an example of the Internet's growing importance to the media. The right URL address can lead to aviation accident reports, campaign contribution databases, people finders, material safety data sheet repositories, tax records for charities and prison inmate tracking systems, to name a few.

The Web for journalists

I used the suggestions for a presentation I gave about how the Internet can help journalists. Here are some of the Web sites we talked about:

Listervs: www.liszt.com/ – The place to locate e-mail discussion groups.

Newsgroups: www.groups.google.com/ – Formerly called Deja News, this newsgroup archive was acquired by the Google search engine and covers just about every topic.

Experts: www.profnetwork.com – A way to find experts online in a variety of fields. Profnet is now requiring journalists to register before they search the database of experts, but registration is free.

Health: www.ncbi.nlm.nih.gov/pubmed/ – Look up abstracts from medical studies.

Courts: <http://pacer.uspci.uscourts.gov>

– Federal court records at your fingertips. Recent cases offer access to the actual court documents.

Lawyers: <http://lawyers.martindale.com/xp/Martindale/home.xml> – Track down lawyers in any part of the country and even outside the United States.

Charities: www.guidestar.org – The nonprofit group Philanthropic Research, Inc. helps donors research other charities. Look up financial information and 990 tax forms that nonprofits file with the IRS.

Type in some keywords, use the control key to select which fields you want to search, and GPO does the work for you.

Worker safety: www.osha.gov/cgi-bin/est/est1 – OSHA offers a large database of company safety violations.

Airline crashes: http://nasdac.faa.gov/asp/fw_ntsb.asp – Using an aircraft's N-number or serial number, you can find out whether a plane has experienced past problems.

Government records: www.access.gpo.gov/su_docs/aces/aaces002.html – The Government Printing Office offers a slow but neat way to conduct keyword searches across multiple federal databases. Type in some keywords, use the control key to select which fields you want to

search, and GPO does the work for you. Includes access to bills and Congressional reports.

Statistics: www.fedstats.gov – An easy-to-use clearinghouse for a variety of federal figures.

Internet: www.netsol.com/cgi-bin/whois/whois – Need to know who's behind an Internet site? Type their Web address and find out.

Politics: www.opensecrets.org – One of many places to look up how much money politicians are raising and who is giving to them.

Open records: www.rcfp.org/tapping/index.cgi – Find out what records are public in any state.

Chemicals: www.msdssearch.com – A repository of material safety data sheets.

Inmates: www.bop.gov/ – Use this site to locate federal prisoners.

The Web by journalists

Some of the best sites on the Internet are maintained by reporters and researchers. Among them:

Investigative Reporters and Editors: www.ire.org – Offers tip sheets, story abstracts and resources to members. Check the listserv archive to see if your topic has been discussed. Also offers a compilation of web sites at www.ire.org/training and breaks down sites by beat at www.ire.org/resourcecenter/initial-search-beat.html.

The Journalist's Toolbox: www.journaliststoolbox.com/

Power Reporting: www.powerreporting.org

A Journalist's Guide to the Internet: <http://reporter.umd.edu>

Assignment Editor: www.assignmenteditor.com/

Fast Facts: <http://gwu.edu/~gprice/handbook.htm>

The Reporter's Desktop:
www.reporter.org/desktop/

Deadline Online:
www.deadlineonline.com

To learn more

Search Engine Watch:
<http://searchenginewatch.com/> —
 Journalist Danny Sullivan reviews
 search engines and offers good tips.
 Sign up for his e-mail newsletter.

The same kind of web site, with e-mail
 newsletters, can be found at Research
 Buzz: www.researchbuzz.com/, by
 Tara Calishain.

Scout Report: [http://scout.cs.wisc.edu/](http://scout.cs.wisc.edu/index.html)
[index.html](http://scout.cs.wisc.edu/index.html) — Researchers at the Univer-
 sity of Wisconsin-Madison seek out au-
 thoritative Web sites in a variety of
 fields and make their recommenda-
 tions available to anyone. Browse their
 archives and subscribe to their free,
 weekly e-mail.

Web Search FAQ: [www.infotoday.com/](http://www.infotoday.com/searcher/oct01/price.htm)
[searcher/oct01/price.htm](http://www.infotoday.com/searcher/oct01/price.htm) — Gary Price
 at George Washington University has
 useful web pages for journalists and
 wrote this article about using search
 engines to your advantage.

Pandia: www.pandia.com/ — Reviews
 search engines, highlights their advan-
 tages and disadvantages.

IRE: www.ire.org — Look up tip sheets
 about mining the World Wide Web for
 information.

Good Internet advice for reporters
 can be found outside the electronic
 realm. *A Journalist's Guide to the*
Internet, by Christopher Callahan,
 and *Find it Online*, by Alan M. Schlein,
 were written by journalists whose
 sites are mentioned above. *Super*
Searchers in the News, by Paula J.
 Hane, is a book of interviews with
 reporters and news researchers
 who talk about how they search the
 World Wide Web.

John Tedesco can be reached by e-mail at
jtedesco@express-news.com

From the Lists

continued from page 13

Looking at test scores

Q: I'm currently helping the edu-
 cation reporter with the state school
 scores released to the media yes-
 terday. I would like to know what
 would be the best approach to cal-
 culate every district's score on
 each category, English, language
 and arts. The same with science.
 ... The education reporter received
 the results on an Excel table. I was
 thinking about basic math: adding
 all the percentages and then divid-
 ing them by the number of schools
 on each district.

A: A better approach might be
 to calculate a WEIGHTED aver-
 age. To do that, multiply the av-
 erage for each school by the
 number of kids who took the
 test, add those numbers to-
 gether, then divide by the total
 number of kids.

Deleting the right records

Q: I'm trying to write a query
 in SQL Server 7.0 to delete
 some data. I have two tables,
 each of which has the same
 unique code field. I want to de-

lete from Table1 all the rows
 where the code matches the
 code field in Table2, but every
 wording I try gives me an error.

A: Try: Delete from table1
 where uniqueid in (select
 uniqueid from table2)

Space, memory problem

Q: I'm trying to change the date
 field from text to a date and it
 worked fine on a smaller
 dataset. Now, on a much larger
 dataset, I get this: "Microsoft
 Access can't change the data
 type. There isn't enough disk
 space or memory." Funny, I
 have more than 30 gigs of free
 space and 256 megs of RAM.
 Seems that would be plenty to
 change a stinking date field.
 What gives? Anyone got any
 ideas? I have workarounds
 aplenty at my disposal, but
 seems like I should be able to
 attack this head on without hav-
 ing to leave Access.

A: Although your database
 may be on a drive with 30 gigs
 of free space, Access may be
 using another drive for the
 TEMP files it creates. To change
 the folder used by Access, go
 to control panels and double
 click on the "System" icon. Go
 to the "Environment" tab, and
 you will see which folders are set
 for TMP and TEMP files. Just
 change the folder settings to a
 folder on the drive with all the
 room and you should be all set.

CONFERENCE

San Francisco 2002

By Gina Bramucci, IRE and NICAR

Reporters and editors are bound for the City by the Bay this year for the IRE National Conference, May 30-June 2. The conference has become an invaluable tradition for newsrooms around the country, offering journalists the chance to enrich their skills through workshops and panel discussions featuring top-notch investigative reporters.

The IRE conference comes at a crucial moment in journalism. With national security high on the public agenda, reporters and editors are facing increasing challenges to freedom of information. The San Francisco conference will present beat basics, while giving special attention to the major stories and information access issues that have emerged in the past year.

Panels will highlight the best investigative journalism from around the nation, delving into energy regulation and deregulation,

immigration, military and defense issues and campaign finance reform. In light of the high-profile collapse of Enron and the Kmart bankruptcy, attendees also can expect tips on using business documents and covering corporate fraud.

This year's conference will integrate international journalists into panels more than ever before. "The issues we cover have clearly become global," said Brant Houston, IRE's executive director. "IRE is expanding and building on the network of journalists helping each other."

With the *San Francisco Chronicle* as a local host, the conference location also will allow for greater focus on coverage of diversity, said Houston. "A great thing about San Francisco is the diversity of the media in scope, ethnicity and approach," he said.

Nuts, bolts and more

In addition, the always-popular presenters Donald Barlett and James Steele will make another appearance this year, sharing what they've learned in more than three decades of prize-winning investigative journalism. (Note their story on the Olympics in the March-April 2002 issue of *The IRE Journal*.)

Fast-track sessions aimed at print, broadcast and online journalists will deliver the nuts and bolts of beat reporting—offering tips for backgrounding individuals and for covering schools, businesses, city halls and courts.

On Thursday, panels dedicated to computer-assisted reporting will help journalists learn and fine-tune their CAR skills. Veteran reporters will collaborate with NICAR staff to tackle issues related to the criminal justice system, as well as campaign finance, the latest census releases, mapping and geographical information systems.

Other panels devoted to computer-as-

readme.txt

Tipsheets

The IRE Resource Center offers a searchable database of tipsheets for journalists at www.ire.org/resourcecenter or call 573-882-3364 to order tipsheets.

sisted reporting will include:

- The converging investigation: TV, print and Web together
- The 20 databases you need to have
- Air security and safety
- Early warning systems: Data on failing infrastructures
- Immigration and labor safety
- Military and defense
- Public health
- Environment and energy

As always, the winners and finalists of the IRE Awards will be announced on Saturday. The awards presentation is an occasion for IRE to pay tribute to the best investigative reporting from 2001.

On Sunday, speakers will expand the work they did last year to help reporters strengthen their skills and build their careers.

Elections, registration

Each year, IRE addresses membership business during the conference. Elections for board of director seats will take place on Saturday, June 1. Members can request an absentee ballot if unable to attend the conference. (See more detail in the March-April 2002 of *The IRE Journal*.) A separate election for IRE Awards contest judges will take place immediately following board elections.

The conference will be held at the San Francisco Hyatt Regency. Members can register online at www.ire.org/training/sanfran/ or photocopy the form found in this issue of *Uplink*. Hotel reservations can be made by calling the Hyatt Regency at 415-788-1234 or 800-233-1234.

Gina Bramucci can be reached by e-mail at gina@ire.org

You're invited

to the Education Writers
Association's
2002 National Seminar:

**"Whose Truths?
Perspectives,
Perceptions &
the Public Good"**

April 25-27, 2002
Washington, D.C.
Wyndham City Center

Explore the ethics of covering testing and accountability, learn about observing classrooms, evaluating research, tackling large projects, changing demographics, & more!
Call 202.452.9830, visit www.ewa.org.



Tech tips...

Automating parsing

By Ron Nixon, IRE and NICAR

A few years back while I was a reporter at the *Roanoke (Va.) Times*, I came across a problem in a database that confounded me for days. I'd been given a database with names in it spelled hundreds of different ways: first, comma, last; last, comma, first; first,

data, trying to write a program to clean and parse names and addresses.

Other programs

One of my favorites programs, NameChop, is available for free from Semaphore Corp.'s Website. Sema-

A parsing program that had gotten attention on the IRE listserv for parsing is ParseRat. The program costs about \$50 and not only parses names, but does addresses as well. According to ParseRat's Web site, it can read most file types and export many file types as well. There is a demo that can be downloaded for 30 days.

Programmers in the open source movement have also developed a parsing and data-cleaning program that could be useful to reporters. It's called Potters Wheel and it not only does data cleaning and name parsing, but statistical analysis as well. The program is available for free.

These programs won't solve all your problems. There will still be databases that you'll come across with dirty data that is not easily cleaned. But these little programs make it easier to get to the thing we love most: analyzing the data and getting to the story.

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

Difficult names to split	sPrefix	sFirst	sMiddle	sLast	sSuffix
D/M Stephen Michael St. John Jr. Ph.D.	D/M	Stephen	Michael	St. John	Jr. Ph.D.
Leonard Bark - Jones ESQ.		Leonard		Bark-Jones	ESQ.
Dr. Michael D. Downer D.D.S.	Dr.	Michael	D.	Downer	D.D.S.
Mr Thomas	Mr			Thomas	
Li. Col. Roger J. Wagner III	Li. Col.	Roger	J.	Wagner	III
mR mARTIN aLISTAIR de La RUE Ph.D.	Mr	Martin	Alistair	de la Rue	PH.D.
GREG MILLER-ADSHHEAD B.A. Dip. Ed.		Greg		Miller-Adshhead	B.A. Dip. Ed.
0000000JOSH STEVENSON JR		Josh		Stevens	Jr
WALKER, SUE @@@@%				Walker	
MRS TIMOTHY JOHN ARTHUR VON BUEREN	Mrs	Timothy	John Arthur	von Bueren	
Abraham, F Murray		F	Murray	Abraham	
Anderson David K		David	K	Anderson	
Freddie "The Mangler" Simonds		Freddie	"The Mangler"	Simonds	
McDoe, ROBERT A		Robert	A	McDoe	
SIR WALTER JOHN RAYMOND BETTS, SNR	Sir	Walter	John Raymond	Betts	Snr
OSCAR DE LA HOYA		Oscar		de la Hoya	
MR C. J. TILMAN - BOAM III	Mr	C.	J.	TILMAN-Boam	III
MR & MRS JACK & JILL MC HILL	Mr & Mrs	Jack And Jill		Mc Hill	
Mr Rob Smith-Jones	Mr	Rob		Smith-Jones	
Mr Art P van der Lay	Mr	Art	P	van der Lay	
MCDOE, MICHAEL S.		Michael	S.	McDoe	
Harry o Gorman		Harry		O'Gorman	
Ms Nina D'Aquino	Ms	Nina		D'Aquino	

space, last; etc. It was a parsing nightmare. After a few days of writing string functions until I was blue in the face, I gave up. Then someone mentioned that there was probably existing software that could help with my problem.

They were right. After searching the Web I found a little program called Splitter made for Microsoft Access databases. The program not only splits names no matter how they are entered into a database, it also guesses the gender and will create a new table leaving your original table alone. The program, made by Australian Terry Chapman, costs about \$40 and is easily downloaded.

Splitter is not alone. There are literally dozens of shareware packages out that that nearly eliminate the need to spend countless hours poring over

phore is a company that cleans mailing lists. NameChop works on .dbf and text files. You create additional fields in your table with first, last, middle initial, suffixes and prefixes and let it rip. The program takes a few seconds and your data problems are gone. In addition to NameChop, Semaphore also gives away DupeDetect, a program that lets you flag duplicate records by selection criteria to match on like first name, last name and Zip Code, etc. The program then flags the duplicates and gives you the option to delete them.

Another program that works fairly well at parsing files is McXtract, made by Pierre R. Geevers, a programmer from Nantes, France. The program allows you to import text files in various formats and parse out information. I've only tried the demo, but it seems to work pretty well with big text files.

readme.txt

Web addresses

Splitter:

www.users.bigpond.com/wemba/splitter.htm

NameChop:

www.semaphorecorp.com

McXtract:

<http://ourworld.compuserve.com/homepages/pierregeeevers/mcx.htm>

ParseRat:

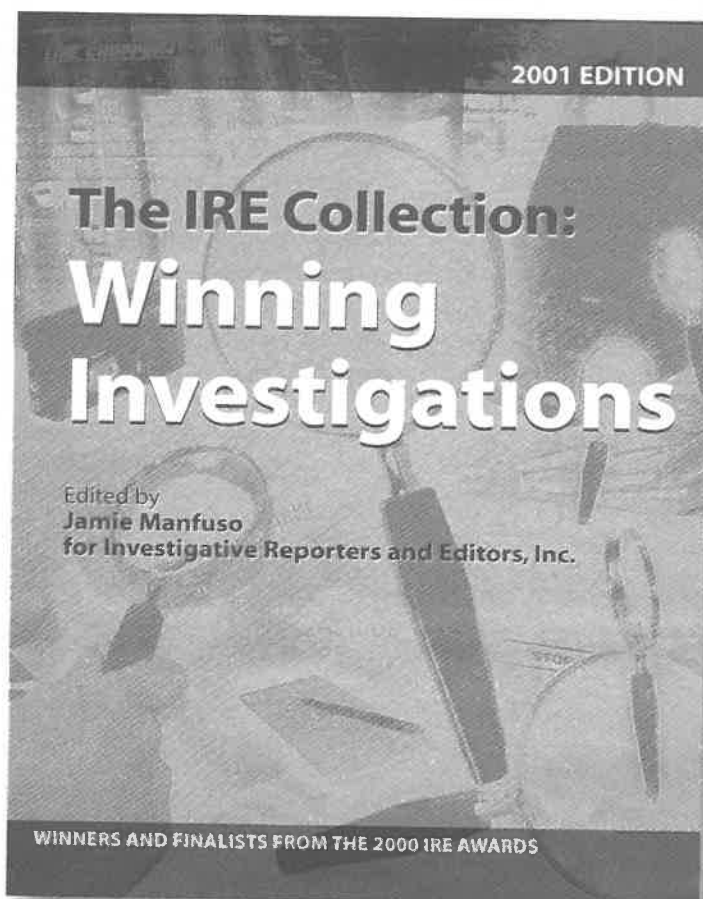
www.guysoftware.com

Potters Wheel:

<http://epoch.cs.berkeley.edu:8000/~rshankar/abc/abc-rel/1.3>

Get the scoop

on the year's
best investigations



The latest compilation book from IRE not only salutes the winners of the prestigious IRE Awards, it offers testimony from the winners.

Information is shared on how the work was completed, which techniques proved most valuable and what resources were utilized to win the print, broadcast, online and other categories. A great guide for launching similar projects in your own back yard!

IRE members: \$16 each • Non-members: \$30 each
Plus first-class postage: \$4 for the first book, \$2 for each additional book

E-mail questions to rescntr@nicar.org

ORDER NOW!

By MAIL

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

By PHONE

Call 573-882-3364
with your VISA or
MasterCard

By WEB

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

CAMPAIGN FINANCE

Money in – and from – the bank

By Aron Pilhofer, IRE and NICAR

Sometimes, the best campaign finance stories happen right in your own back yard. That's what the *News Journal* in Wilmington, Del., discovered when the paper investigated a sudden upsurge in campaign spending by the state's second largest employer: MBNA, one of the nation's largest issuers of credit cards.

Reporter Robert Long, who led the project, became interested after noticing an unusually large number of contributions coming from company executives.

"After seeing MBNA ranking high in some presidential races, we started to take an in-depth look at their campaign activities," he said.

Poring through FEC reports in the spring of 2000, when the paper first began its examination, the *News Journal* found that MBNA executives were the single largest identifiable group of contributors to then-candidate George W. Bush's presidential campaign, giving more than \$250,000.

At the time, MBNA was second only to investment bank Goldman Sachs in total contributions to presidential candidates.

Following the money

In several follow-up articles, the *News Journal* documented the company's lavish campaign spending on state and national candidates and party committees throughout the 2000 election cycle.

MBNA ended up as the most generous campaign contributor among finance/credit companies in 2000, giving more than \$3.5 million to federal candidates and committees. That's

\$1 million more than the company contributed during the previous four election cycles combined, according to the Center for Responsive Politics.

In the course of reporting the story, the *News Journal* uncovered memos that showed how top MBNA executives were influencing employees to contribute to candidates the company favored.

Company executives said the firm was interested in promoting good government. But campaign-finance reform advocates said it was more than coincidental that the company's interest in national politics came at a time when an important piece of bankruptcy reform legislation was pending in Congress.

MBNA and others in the credit industry had been pushing hard to make it

TOP RECIPIENTS

Below are the politicians who have received \$10,000 or more during the 1999-2000 campaign season from MBNA employees and their households. This data came from computerized and paper campaign finance records filed with federal and state election offices.

Recipient	Amount	Party	State
Bush, George W	\$243,250.00	Rep	TX
Burris, John	\$147,400.00	Rep	DE
Ruppersberger, Dutch	\$107,850.00	Dem	MD
Roth, William V Jr	\$99,501.00	Rep	DE
Kasich, John R	\$94,250.00	Rep	OH
Biden, Joseph R Jr	\$62,200.00	Dem	DE
Kyl, Jon L	\$27,000.00	Rep	AZ
Hutchison, Kay Bailey	\$13,000.00	Rep	TX
Baker, Richard	\$10,000.00	Rep	LA
McCollum, Bill	\$10,000.00	Rep	FL
LaFalce, John	\$10,000.00	Rep	NY
Lott, Trent	\$10,000.00	Rep	MS

readme.txt

Story

Read the MBNA story at: www.delawareonline.com/newsjournal/archives/mbna/0312200a.html

more difficult for those who file bankruptcy claims to walk away from debts they owe. Following the election, both houses of Congress passed similar bankruptcy reform bills.

The paper also tracked MBNA's soft money donations and donations made to all of the members of congressional banking committees — and found a close correlation between whom the company was giving to and how committee members voted on bankruptcy legislation.

Compiling the data

The project involved compiling data from several campaign finance sources:

- To compile the federal numbers, the *News Journal* used raw FEC data, downloaded from the agency's Web site.
- Delaware had not yet begun computerizing state filings, so that data had to be typed into a database by hand.
- The paper also keyed in some data from paper filings submitted by MBNA's federal political action committee.

As with any campaign finance project, there were problems.

"In many of the FEC filings, MBNA was listed as an employer for MBNA employees. But in the state reports and some of the FEC reports, they weren't," Long said. "So we built a database of corporate officers — which we got from annual reports, PAC reports and other sources — compiled addresses from voting records and were able to use those to match

continued on page 22

money

continued from page 21

spouses who were also donating to Bush and other Republicans."

Fixing the problems

Other problems included:

- Finding executives' spouses, which involved some guesswork because the FEC does not include the full address of all contributors. The *News Journal* used a methodology similar to that used by the Center for Responsive Politics: Persons with the same last name; listing the same city, state and ZIP Code; contributing the same amount to the same candidate on the same day were considered related.

- Juggling many different databases, from different data sources, which made the analysis tricky.

- Managing a moving target. The paper constantly had to update numbers as the company made more contributions, and the reporters found more ways in which MBNA was contributing.

All in all, it wasn't a pretty collection of data, Long said, "but it was accurate and got the job done."

Aron Pilhofer, formerly with the *News Journal*, can be reached by e-mail at aron@ire.org

CAMPAIGN FINANCE

Students reveal gifts to legislators

By Jeff C. South, Virginia Commonwealth University

In computer-assisted reporting, there's a tendency to think that bigger is better: that the more records you crunch, the more impact your story will have.

But my Legislative Reporting students at Virginia Commonwealth University dispelled that notion last spring when they created and analyzed a relatively small database: a list of all gifts received last year by state legislators, the governor and other top officials.

The dataset contained fewer than 1,000 records, totaling about \$185,000 in gifts, taken from the conflict-of-interest forms officials filed in January. However, the resulting stories created quite a buzz at the Capitol and, thanks to AP, resonated across the state.

We put our stories online, with a searchable database detailing the gifts reported by each legislator. We highlighted the biggest gifts – including trips to hunt caribou in the Arctic Circle and quail in Georgia. And we raised ethical questions: Should lawmakers be accepting such freebies from businesses, interest groups and lobbyists seeking

to influence the legislative process?

Story can get response

The gifts project struck a louder chord than I had expected. Stateline.org, a Web site devoted to state government issues, wrote a story about our efforts, and several journalists praised our efforts. Our Web site received thousands of hits – and numerous visitors sent appreciative e-mails. We got more response on our gifts package than on a much larger CAR project my spring 2000 class did on campaign contributions.

Because that project was so time-consuming, I planned to scale back this year. I figured the gifts project would be more manageable: We'd get the Statements of Economic Interests filed by each legislator, type into Excel the information about gifts, then sort and subtotal the data.

Not terribly flashy, I thought, but the project will teach students about the importance of getting – and computerizing – public records.

Drawing attention

In retrospect, here are some reasons

ICIJ Award for Outstanding International Investigative Reporting

INTERNATIONAL CONSORTIUM



OF INVESTIGATIVE JOURNALISTS

A PROJECT OF THE CENTER FOR PUBLIC INTEGRITY

A \$20,000 first-place prize and up to five \$1,000 finalist awards aim to recognize, reward, and foster international investigative reporting.

Journalists of any nationality working in print, broadcast and online media are eligible to apply. Books are not accepted. The work must have involved reporting in at least two countries and must have been first published or broadcast in general information media between June 1, 2001 and June 1, 2002. Deadline to apply is July 15, 2002.

For more information on the ICIJ Award, as well as the 2001 winners and finalists, see www.icij.org or call 202-466-1300.

why our package on legislative gifts drew attention – and why you might want to put such a story on your to-do list:

- Many gifts have a cachet that campaign donations don't. A lobbyist may get more ingratiating bang for the buck with a \$100

limits on the amount or size of gifts or campaign donations legislators can receive.)

- Campaign donations often are online (in Virginia, thanks to the Virginia Public Access Project [www.vpap.org]). Not so for gifts; those records had been avail-

Search the legislative gifts database

VCU's Legislative Reporting class compiled a database of gifts that state legislators reported receiving during the year 2000. The information came from the Statements of Economic Interests filed in January by each member of the House and Senate.

Legislator's last name:

District	Legislator	Giver's name	Description of gift	Value
H70	Jones, Dwight Clinton	Philip Morris	Holiday gift box	\$113.00
	Jones, Dwight Clinton Total			\$113.00
H76	Jones, S. Chris	Glaxo Wellcome	dinner	\$96.18
H76	Jones, S. Chris	Natl. Ass. of Chain Drug Stores	state issues conference, Palm Springs, CA	\$1,675.00
H76	Jones, S. Chris	Ringling Bros. Circus	4 circus tickets	\$68.00

steak dinner than with a \$100 or even \$1,000 check to a campaign committee. In Virginia, where \$5,000 campaign contributions are pretty common, an interest group could spend less money but make a bigger impression by sending a key legislator on a vacation.

- Gifts often guarantee face time with lawmakers – giving lobbyists an opportunity to state their case for or against legislation. We found that legislators received \$26,000 in meals and \$24,000 in trips, with lobbyists usually serving as hosts. Access is the goal of most big gifts and campaign contributions, according to government watchdog groups.

- Constituents may view certain gifts more suspiciously than campaign contributions. People might sympathize with politicians' need for money to run for office. But they might think it's grubby for Virginia lawmakers to accept free Washington Redskins tickets, golf equipment and eyeglasses – things ordinary folks must pay for.

- The rules governing gifts may be looser than the rules on campaign donations. In Virginia, lawmakers cannot accept campaign contributions during the legislative session – but they can accept gifts any time. (Virginia has no

able only on paper in Richmond. Our project made the gifts data available to the public for the first time over the Internet.

How the work progressed

If you're interested in doing a similar project in your state, first get the rules on legislative gift-giving. You can find them on the Web site of the National Conference of State Legislatures (<http://www.ncsl.org/>).

Here's how we did our investigation:

At the start of the session, we ordered copies of the Statements of Economic Interests for members of the House of Delegates (from the House Clerk's Office) and for members of the Virginia Senate (from the Senate Clerk's Office). These statements, also called conflict-of-interest forms, list legislators' debts, investments, employers, real estate holdings and other financial information.

We zeroed in on Schedule E, on which legislators must list all gifts of \$50 or more received the previous year. We had 140 legislators to contend with (100 House members and 40 senators); the class was made up of 13 students and an instructor. So we each took the records for 10 legislators and entered the Schedule E information into a spreadsheet.

We did the data entry in Microsoft Excel. For each gift, students entered the recipient's name, the giver's name, the giver's city and state, the description of the gift, the value of the gift and any notes made on the legislator's form.

We then combined each student's data into a single spreadsheet – with all gifts on all legislators. We fixed typos and inconsistencies, especially in the spelling of donors' names. Then we analyzed the data with Excel's PivotTable Report feature. This allowed us to see how much each donor had given in gifts and how much each lawmaker had received.

Stories flow from data

After doing the analysis, the class discussed possible stories: about the top givers and recipients, about expensive or unusual gifts, about why donors give gifts to legislators.

One student decided to write about gifts given to Gov. Jim Gilmore, Lt. Gov. John Hager (who presides over the Senate) and Attorney General Mark Earley. These officials file their Statements of Economic Interests with the Secretary of the Commonwealth.

After obtaining those forms, the class again divvied up the data entry so we could sort and total the top officials' gifts in Excel. (It would have been hard to do this by hand or with a calculator: Gilmore reported 337 gifts worth almost \$55,500.)

To guide us in this project, we found an article Kit Wagar of *The Kansas City Star* did for the September 1999 issue of *Uplink*. Using Excel and FoxPro, Kit had analyzed the gifts given to the Missouri Legislature.

Jeff South can be reached by e-mail at jcsouth@vcu.edu

readme.txt

Story

The URL for the package of stories on legislative gifts is: www.people.vcu.edu/~jcsouth/on-the-lege/

workplace

continued from page 1

& NICAR in this issue offers some guidance for journalists who want to stay afloat while working with the OSHA data.

These articles just touch on what you do. There are plenty of other angles that data-savvy reporters can follow.

Pensions

After the fall of Enron flattened out its employee retirement savings plans, heavy in company stock, St. Paul Pioneer Press business reporter Jim McCartney turned to the web for the local angle.

McCartney wanted to see whether the retirement savings plans offered by the 15 biggest publicly-traded companies in the Twin Cities were loaded with company stock. So he turned to the

crunching? You can obtain the entire ERISA database from the U.S. Department of Labor, and compare some 750,000 pension plans across the country. Pension plan administrators

Form 5500
Department of the Treasury
Internal Revenue Service
Department of Labor
Pension and Welfare Benefits Administration
Personnel Benefits Coverage Corporation

Annual Return/Report of Employee Benefit Plan
(With 100 or more participants)
This form is required to be filed under sections 104 and 4085 of the Employee Retirement Income Security Act of 1974 and sections 6038D, 6047(b), 6057(b), and 6058(a) of the Internal Revenue Code, referred to as the Code.
See separate instructions.

OMB Nos. 1510-0016
1216-0009
1998
This Form is Open to Public Inspection.

For the calendar plan year 1998 or fiscal plan year beginning January 01, 1998, and ending December, 1998

If A(7) through A(4), B, C, and/or D, do not apply to this year's return/report, leave the boxes unmarked.

A This return/report is: (1) ☐ the first return/report filed for the plan; (2) ☐ an amended return/report; (3) ☐ the latest return/report filed for the plan; or (4) ☐ a short plan year return/report (less than 12 months).

IF ANY INFORMATION ON A PREPRINTED PAGE 1 IS INCORRECT, CORRECT IT. IF ANY INFORMATION IS MISSING, ADD IT. PLEASE USE RED INK WHEN MAKING THESE CHANGES AND INCLUDE THE PREPRINTED PAGE 1 WITH YOUR COMPLETED RETURN/REPORT.

B Check here if any information reported in 1a, 2a, 2b, or 3a changed since the last return/report for this plan. ☐

C If your plan year changed since the last return/report, check here. ☐

D If you filed for an extension of time to file this return/report, check here and attach a copy of the approved extension. ☒

1a Name and address of plan sponsor (employer, if for a single-employer plan)
(Address should include room or suite no.)
ENRON CORP
1400 SMITH ST
HOUSTON, TX 77002-7327

1b Employer identification number (EIN)
47-0255148

1c Sponsor's telephone number
713-853-8161

1d Business code (see instructions, page 20)
2212

1e CUSIP issuer number
293561

readme.txt

Resources

Tipsheets and stories from the IRE Resource Center can assist journalists working on similar stories.

Tipsheet #935

A brief overview of laws affecting workplace reporting, reviews some of the agencies that oversee employers and a listing of sources valuable to writing and producing stories about work. The packet includes a listing of Web sites designed to introduce you to government databases on workplace-related information.

To order, call (573) 882-3364 or go to www.ire.org/resourcecenter.

Web addresses

Department of Labor:
www.dol.gov/dol/pwba/public/pubs/main.htm

10k Wizard:
www.10kwizard.com

FreeERISA.com:
www.freeERISA.com

online 10k Wizard and looked for 11K annual reports the companies submitted to the U.S. Securities Exchange Commission that provide detailed financial information about employer 401(k) retirements savings plans. Not all plans are covered; only those in which the employer matches employee contributions in company stock, and grants the employees the right to buy more shares. McCartney said he found the data on 10k Wizard, a subscription service, up to date.

McCartney obtained data about other pension plans at FreeERISA, a site that offers no-cost access to pension plan data collected by the Department of Labor's Pension and Welfare Benefits Administration under the Employee Retirement Income Security Act. McCartney said that the data for some companies was outdated by as much as three years.

After looking over the documents and punching the numbers into his calculator, McCartney wrote a Feb. 2 story saying two-thirds of the top companies had 401(k) plans in which company stock accounted for at least 40 percent of the total assets. McCartney found two companies based in the Twin Cities – Target and Best Buy – with company stock concentrations higher than Enron's.

Comfortable with heavy-duty number

every year submit detailed information about the plans using Form 5500, available with other documents on the Department of Labor's Web site. In the coming months, IRE and NICAR will be selling the pension data in an easy-to-use format.

Five years ago, News 12 Long Island reporter Mark Lagerkvist obtained the massive database (a whopping 6gb on 27 9-track tapes!) from the department via a federal Freedom of Information Act request. After carving out smaller data sets and analyzing them, Lagerkvist broadcast "Stolen Dreams," a special report on how employers plundered pension plans.

Foreign/immigrant workers

Soon after the Sept. 11 attacks, federal agents began sweeping the United States for suspected terrorists, focusing on Middle Eastern men. Foreign visitors, legally in the United States with temporary work visas, were questioned. So were legal immigrants.

So where are these workers? Data from the Immigration and Naturalization Service, part of the U.S. Justice Department, can provide some clues. The INS data, sold by the NICAR database library, specifies the sex, age, marital status, country of birth, coun-

try of last permanent residency, port of entry and ZIP Code of intended residence for immigrants.

With some rudimentary queries in your database manager, you'd be able to find out what national groups live in your area, and see which occupations are the most prevalent for the groups.

Workers compensation

State workers compensation data, used by journalists investigating bogus injury claims and injury patterns at problem work sites, can be put to other uses.

Last year, reporters at the Washington Post dug through ten years of Maryland Workers Compensation Commission data as they probed the Prince George's County Police Department. [For details, see "Detailing Questionable Police Shootings Required Extensive Search of Documents" in the March-April 2002 issue of *The IRE Journal*.] While examining fatal shootings by police, the reporters mined the data for cases where police officers filed disability claims because of the stress and emotional trauma related to the shootings. One field in the database covering 700,000 unique claims sometimes contained a detailed narrative of the shootings. After identifying the cases, the reporters obtained the original case documents on file with the commission.

Post reporters used the data again when they were looking at out-of-control police dogs. This time, the reporters queried for cases where police filed disability claims because they were bitten by department canines.

Other journalists have put workers compensation data to use in more traditional ways. In 1997, Mary Fricker and Bob Norberg, reporters at the *Santa Rosa Press Democrat* analyzed workers comp data and unmasked claims that fraud was behind one third of the disputed cases.

David Herzog, formerly an investigative reporter with *The Providence Journal*, can be reached by e-mail at dherzog@nicar.org

Upcoming CAR Training

Advanced Statistics Workshop, June 9-13 – Chapel Hill, N.C.

CAR guru and Knight Chair in Journalism Professor Philip Meyer once again will present an advanced statistics workshop with IRE and NICAR. The workshop strengthens the skills of reporters who want to move beyond basic computer-assisted reporting and use social research methods and statistical analysis in their work, including understanding and using regressions, indexes, sampling and polling. Participants are introduced to SPSS, a statistical software package. Other instructors include Sarah Cohen of The Washington Post and IRE Executive Director Brant Houston.

Mapping the Census Seminar, May 3-5 – Columbia, Mo.

IRE and NICAR will offer this intensive hands-on training using ArcView GIS software, with a focus on mapping Census data to uncover stories. We will look at noteworthy stories that have used mapping. The sessions will include address geocoding, thematic mapping, overlaying and spatial analysis. Trainers include Andy Lehren of Dateline NBC, Jennifer LaFleur of the St. Louis Post-Dispatch and David Herzog of NICAR and the Missouri School of Journalism. Participants are encouraged to bring their own data to work on during open lab time.

Using the Religious Census, April 10-11, 2002 – Washington, D.C.

Religion Newswriters Association Workshop at National Press Club: Reporters will learn how to use data from Glenmary Research Institute's religious census, which will be released in May or early June. The Census is done every 10 years by the Institute, a religious-based research organization. The Census Bureau often refers reporters to this organization since they are barred from asking religious questions. The religious census contains information relating to the number of churches and total church membership by denomination by state and county.

CAR Boot Camps

• March 24-29 and May 19-24 – Columbia, Mo.

IRE and NICAR are offering two CAR boot camps in Missouri this spring with special tracks. The six-day boot camps each open with special sessions on campaign finance and each offer four fellowships for journalists at small- to medium-sized news organizations who are interested in improving their coverage of elections. The May boot camp also will have special sessions on using Census data.

• Aug. 4-9 – Columbia, Mo.

A special track for broadcast journalists will be featured at this Boot Camp. Learn how to acquire electronic information, use spreadsheets and databases to analyze the information and to translate that information into high-impact stories. In addition, NICAR provides follow-up help when participants return to their news organizations.

For registration and fellowship information, go to

www.ire.org/training/otr.html

Uplink Info

A newsletter of the National Institute for Computer-Assisted Reporting

Editor

Brant Houston
brant@ire.org

Managing Editor

Jeff Porter
jeff@nicar.org

Asst. Managing Editor

David Herzog
dherzog@nicar.org

Contributing Editors

Mike Sherry
Amy Sherrill

Art Director

Lisa Triefenbach
lisa@ire.org

Editorial Phone

573-884-7711

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

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

Director of Publications

Len Bruzzese
len@ire.org

Advertising Coordinator

Pia Christensen
pia@ire.org

Subscription Administrator

John Green
jgreen@ire.org

Subscriptions

IRE members \$40, nonmembers \$60

Uplink Address:

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

*Postmaster: Please send
address changes to IRE-NICAR.*

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

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

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