

# Uplink

May 1999

A newsletter of the National Institute for Computer-Assisted Reporting

## MAPPING ANALYSIS

### Charting a flood

By Chris Williams  
*San Antonio Express-News*

Two weeks after a freak storm dropped more than 15 inches of rain on San Antonio, more than 100 residents from a flood-ravaged neighborhood gathered in an elementary school gym. They wanted to know why the government, at any level, had not told them they were at risk of flooding.

Across town, journalists at the *San Antonio Express-News* already were working on the same deceptively simple questions. Before the swollen creeks had receded, we were asking: "Where did the flood water go? Was there any

pattern to the damage? Did people who needed flood insurance have it? Why not?"

Some moderately sophisticated CAR techniques were needed to answer those simple questions. On Nov. 29, five weeks after the flood, we printed the results of the work of a computer-savvy reporter and news researcher. We found:

- The path of the destruction closely followed the 100-year flood plains published in 1996. In fact, 75 percent of the 175 residences destroyed by the flood were in those flood plains.

- Citywide, only about 10 percent of the damaged buildings were insured by the National Flood Insurance Program, which meant 90 percent of the people in 470 damaged buildings were trying to rebuild on slim federal grants or disaster loans they could not afford to repay.

- Many more houses along creeks on the poor side of town were damaged than buildings along creeks in wealthier areas.

#### Geocoding databases

The story demonstrated that you don't need a full-time data analyst or CAR guru to get high-impact stories into the newspaper.

From the very beginning, we knew that we wanted to print a map showing the locations of hundreds of properties damaged by the flooding. We quickly decided the way to do that was building, then geocoding, a database of damaged addresses. Geocoding is the geek word for assigning latitude and longitude coordinates to (in our case) street addresses.

We used the software at hand — Microsoft Access for building the database and MapInfo for generating the maps and analyzing the relationship between the flood plains and the locations of the actual damage.

The American Red Cross gave us hard

Continued on page four

### Inside Uplink

We at NICAR would like to dedicate this issue of *Uplink* to its managing editor, Brent Johnson, who died April 13 after complications from heart surgery. He was 27. Editing many of these stories from his hospital bed, Brent kept working until the night before the surgery. We'll miss his dedication. Please see page 2.

This issue of *Uplink* includes reports of data analysis from across the Atlantic and south of the Rio Grande. Pedro Enrique Armendares of Periodistas de Investigación, a Mexican affiliate of IRE, writes about CAR activity throughout Latin America. Flemming Svith and Nils Mulvad describe the birth of the Danish Institute for Computer-Assisted Reporting. And Dick van Eijk shows how he broke new ground in the Netherlands with his 1994 mapping of election results.

Closer to home, Chris Williams tracks flood-plain data, and Laura Zuckerman discovers a rising boat-accident rate.

PAGE 4

On the Internet

PAGE 12

Handout of the Month

#### CAR REPORT FROM LATIN AMERICA

### Mexico and beyond

By Pedro Enrique Armendares  
*Periodistas de Investigación*

Training Mexican journalists in the methods of investigative reporting has been the purpose of Periodistas de Investigación for almost three years now. PI was the brainchild of IRE members from both sides of the border and it extends IRE's programs, such as workshops on computer-assisted reporting, to Mexico thanks to a grant from the McCormick Tribune Foundation.

In September 1996 Lise Olsen and I opened PI's office in Mexico City. Lise came out of the United States with many years of experience in investigative journalism and CAR. She had also taught CAR workshops for IRE. I began my work in radio and television, later moving to the Mexico City daily *La Jornada*, where I covered special assignments for seven years.

The last decade has been an exciting time for investigative journalism in

Continued on page five

# Uplink

May 1999

Volume 11, Number 4

A newsletter of the National  
Institute for Computer-Assisted  
Reporting

## EDITOR

Brant Houston  
brant@nicar.org

## MANAGING EDITOR

Brent Johnson

## SENIOR CONTRIBUTING EDITORS

Len Bruzzese

len@ire.org

Anna Brutzman

anna@nicar.org

Cindy Eberling

cindy@ire.org

Richard Mullins

richard@nicar.org

## COPY EDITOR

Jeanine M. Davis

## ART DIRECTOR

Wendy Charron

## DESIGNER

Kerrie Kirtland

## STAFF

Britton Clapp

Dawn Fallik

Jason Grotto

Ben Lesser

Ted Peterson

Noemi Ramirez

Aaron Springer

Neil Warner

*Uplink* is published every month  
by the National Institute for  
Computer-Assisted Reporting,  
138 Neff Hall Annex  
Columbia, MO 65211.

(573) 882-0684. Subscriptions are  
\$40 for IRE members, \$60 for  
nonmembers.

Postmaster: Please send address  
changes to NICAR.

Send email to [jeanine@nicar.org](mailto:jeanine@nicar.org)

NICAR is a joint effort of  
Investigative Reporters and  
Editors and the University of  
Missouri School of Journalism.  
NICAR services include hands-on  
newsroom training in computer-  
assisted reporting, special  
academic and advanced training  
in data analysis.

NICAR is supported by grants  
from The Freedom Forum and  
other foundations intended to  
help the institute deliver its  
services nationwide to news  
organizations and associations.

# In Memorium

Many of you knew Brent Johnson, a graduate student here at Missouri, who has been managing editor of our newsletter *Uplink* and has contributed to the *IRE Journal*. Brent died April 13 of cardiac arrest.

Brent had Marfans Syndrome, a connective-tissue disorder that affects the cardiovascular and skeletal systems, and had gone into the hospital to have a heart valve replaced.

Although this was a serious operation, he (and we) fully expected a quick recovery and for him to go on to his internship at the *Miami Herald* this summer. However, complications ensued during and after the operation.

Brent, who was with us for the past two years,



BRENT PAUL JOHNSON  
1971 - 1999

was a gifted writer and reporter and a wonderful colleague and friend with a great sense of humor and perspective. We will miss his voice.

In his memory, we are setting up at IRE a scholarship fund in his name to be given each year to the graduate student who is managing editor of *Uplink*. If you want to make a contribution, please make out your check to IRE and mark it for the "Brent Johnson Scholarship Fund."

## TIME, SWEAT AND SHOE LEATHER

# CAR's advantages

By Paul Overberg  
*USA Today*

One of computer-assisted reporting's less-celebrated virtues is its ability to save time, sweat and shoe leather.

When we sell CAR to skeptical editors and reporters, we often focus on how it can root out corruption or do the previously impossible, such as analyzing and publishing to the Web a school test-score package.

But it's good to remember CAR's humbler virtues. Time and money are always short. Saving you two days of interviewing and saving the editor's plan to run your story the following weekend win just as many converts as a big hit.

This idea struck me again last year when discussion arose on the NICAR mailing list about how one might track and analyze the tragedy of fatally abused children who have been written off as victims of sudden-infant death syndrome. I watched colleagues at Gannett News Service do this in 1991. Most cases turned up from interviews and Lexis-Nexis searches. But they also had a statistician analyze 49,569 death certificates from 1987 of children under nine to determine autopsy rates for various causes of death, especially suspicious ones, in every state. That steered them to good places for shoe-leather work, which turned up great material for the case database.

Reporters Marjie Lundstrom and Rochelle Sharpe won the Pulitzer.

## More recent examples:

- In analyzing county-level personal bankruptcy rates in 1997, *USA Today's* Barbara Pearson checked several correlations before finding that one of the best was whether a state allowed wage garnishment. She mapped the data and found the Augusta, Ga., metropolitan area offered a stark contrast. Flanking the Savannah River, Richmond County, Ga., and Aiken County, S.C., were pretty similar, but the former had twice the bankruptcy rate. Reporter Tanya Albert went to Richmond County and found a wealth of local lore—the perfect prism for this CAR story.

- Water well tests around Oakland County, Mich., last year showed fairly high levels of arsenic. The *Detroit Free-Press* mapped test results with geological information. The worst results clustered where soil was shallow over sandstone from which the arsenic came, said Heather Newman, a *Free-Press* projects reporter and CAR specialist. That told reporters exactly where to find people who were most affected.

- In Utica, N.Y., gasoline prices well above the national and regional averages prompted an editor at *The Observer-Dispatch* to wonder if they were due to concentrated gas station

**Continued on page thirteen**

FOR MORE FLOOD  
RESOURCES SEE IRE  
RESOURCE CENTER  
TIPSHEET #186. IT'S THE  
ST. LOUIS POST-  
DISPATCH'S  
"BIBLIOGRAPHY OF FLOOD  
SOURCES" LISTING  
GOVERNMENT AND NON-  
GOVERNMENT SOURCES  
USED TO REPORT THE  
FAILURE OF NATIONAL  
FLOOD PROGRAMS.  
CONTACT IRE'S RESOURCE  
CENTER AT (573)882-3364  
TO GET IT.

CHECK OUT THE FEDERAL  
EMERGENCY MANAGEMENT  
AGENCY'S WEB SITE -  
FEMA.GOV - FOR INFO ON  
FLOOD SAFETY, FLOOD  
INSURANCE, FLOOD LAWS  
AND REGULATIONS

From page one:

## Flood analysis

copies of their damage assessments, which included addresses, an inspector's opinion of the extent of damage and the estimated depth of the water. We built the database using a form in Access. Although most of the fields were standard, we devised a clever trick that we recommend for anyone who needs to build and then geocode a database.

We copied the streets file out of our MapInfo software and imported it as a table into Access. We cut out all the columns except the street names. Then we deleted duplicate names, which whittled the table down from 87,000 to 20,000 rows. Within our data-entry form, we linked the street address field via a combo box to our new table of street names. The computer would recommend the proper spelling of a street name after we typed its first two or three letters, which saved time and ensured that the street names in the database would be exactly as they appeared in the MapInfo file.

Having street names in a database that don't match those on the streets file of the GIS software is frustrating. The computer can't tell the difference between "Villa Main" and "Villamain." When we were done, we could generate a map of red dots for the 470 buildings damaged in the flood. With MapInfo's query ability, we could sort out homes, businesses, apartments and mobile homes.

### Art and interviews

The missing piece was the flood-plain data. That was easy. We bought it. The MapInfo Web page listed a company, On Target Mapping, that sold digital maps from the Federal Emergency Management Agency for about \$200. From there, we printed a large map of the city and several smaller, blow-up maps of the damaged areas for the art department. We had difficulty outputting maps into a format our art department could import on their Macs. So, the artists scanned the images into their graphics software, then retouched them.

Once the data wrangling was finished, the story became a matter of traditional reporting. We interviewed dozens of people, such as the mother who lost everything in the flood except five dresses and two school uniforms for her son and the city public works director who claimed he told several unheeding neighborhood groups about the dangers of flooding.

After the story appeared, several members of the San Antonio City Council pledged to force the city to notify by mail everyone who lived in a flood plain about their risks of flooding. The council also embarked on a multi-million-dollar program of buying out properties in the flood plain that were damaged by the flood.

Chris Williams can be reached by e-mail at [cwilliams@express-news.net](mailto:cwilliams@express-news.net)

ON THE INTERNET

## Calculating Aids

By Stephanie Reitz  
The Hartford Courant

1. Statistics Every Writer Should Know:  
[nilesonline.com/stats/](http://nilesonline.com/stats/)

Compiled by Rocky Mountain News webmaster Robert Niles, this site offers a great primer on averages, mean, median, per capita and other concepts.

2. The 10 Commandments of Polling:  
[metalab.unc.edu/journalism/pollcmd.txt](http://metalab.unc.edu/journalism/pollcmd.txt)

Ken Blake of University of North Carolina (chapel Hill) explains the basic dos and don'ts of polling. Thou shalt not miss this site.

3. The Spreadsheet Page: [j-walk.com/ss/](http://j-walk.com/ss/)  
Compiled by a computer consultant in San Diego, the best offerings on this site can be

found along the toolbar on the left-hand side of the page. Look for easy tips (and jokes) on various kinds of spreadsheet programs.

4. Martindale's Calculators Online: [www-sci.lib.uci.edu/HSG/RefCalculators.html](http://www-sci.lib.uci.edu/HSG/RefCalculators.html)

It takes a bit of navigating to get used to the format of this page, but the results are worth the effort. It claims to have more than 8,200 calculators on the site, ranging from basic percent-change calculators to translation of names from English into various languages (not helpful during budget time but fun after deadline).

Stephanie Reitz can be reached by e-mail at [reitz@courant.com](mailto:reitz@courant.com), or by phone at (860) 647-5349.

# Title IX undermined

By Jim Hopkins  
The Courier-Journal

In Fourmile, on the western edge of Appalachian Kentucky, softball coach Rick Nelson bristled at the way some high school girls were treated by the Bell County School District.

At the girls' softball field, about half the lights didn't work. Dugouts were little more than tin roofs propped up by two-by-fours. The restrooms were crowded, dirty and didn't always work. Seven miles down the road the boys' baseball team had manicured fields that were amply lit and featured cinder-block dugouts painted with the Bobcats' logo.

Nelson told *The Courier-Journal* that he believed the county school district was in violation of Title IX, part of a 1972 federal law that requires schools to provide comparable athletic programs for girls and boys.

That interview prompted a team of *Courier-Journal* reporters to conduct a six-month computer-assisted examination of athletic programs at hundreds of public high schools in Kentucky and Southern Indiana. Our conclusion in February 1998: There was rampant gender discrimination in equipment, practice times, facilities, budgets and the number of teams offered.

*The Courier-Journal's* disclosures brought immediate results. The Kentucky Board of Education now requires all high schools to certify annually that they're complying with Title IX. For the first time, the state sponsored two Title IX training sessions for school officials. And in February of this year, after the newspaper published a follow-up report showing continued inequities, education board members accused the Kentucky High School Athletics Association of not working hard enough to deal with the problem.

## How we did it

The newspaper's findings came after lots of traditional shoe-leather journalism mixed with newer computer-assisted reporting and analysis techniques.

Under open-records laws, we mailed requests to each of the 255 Kentucky and Southern Indiana high schools that have interscholastic sports. We asked for sports participation numbers by gender, for spending on boys' and girls' sports, for a gender

breakdown on athletic directors and coaches and other information. We also obtained electronic copies of schedules for more than 5,000 boys' and girls' basketball games for the 1997-98 season.

Using Paradox and Excel, we created dozens of databases to track the following: spending, growth in boys' teams versus girls' teams and trends in the hiring of women. We showed, for example, that:

- Although boys comprised about 50 percent of school enrollment in Kentucky's public high schools, they held 62 percent of all the team positions.
- Men held virtually all the coaching slots in major boys' sports and the majority of slots in girls' softball, basketball and soccer.
- Nearly 40 percent of boys' basketball games were played during "prime time" — Friday nights, when audiences are especially large. Less than 13 percent of girls' games were scheduled during that time slot.

## Problems we overcame

The biggest obstacles we faced were in gathering data. Because there was no central repository for the records we needed, the lead reporter had to write individual records requests to each school. When school officials didn't respond, we placed follow-up phone calls and made additional written requests.

We were also hampered by the fact that school officials generally provided raw data, which meant sorting through copies of budgets and hiring schedules to learn, for example, how much was spent on travel for a football team or how many women were in coaching positions.

Finally, to keep the story focused on the people most affected — the students — we crisscrossed the state to interview parents, team members and school officials, then supplemented that reporting by talking with experts on a national level.

Going beyond the data led us to people like Jerry Owens, principal at Owen County High. "There is probably not a high school in the state that isn't violating the law," he said. "The good ol' boy network is very much in place."

Jim Hopkins can be reached by e-mail at [hopkinsj@courier-journal.com](mailto:hopkinsj@courier-journal.com)

SEE THE REQUIREMENTS  
MANDATED BY TITLE IX OF  
THE 1972 EDUCATION  
AMENDMENTS:  
[ED.GOV/OFFICES/OCR/  
INTERATHL.HTML](http://ED.GOV/OFFICES/OCR/INTERATHL.HTML)  
• IN 1997, THE KANSAS CITY  
STAR PUBLISHED A SIX-PART  
PIECE ON THE NCAA THAT  
INCLUDED STORIES AND  
DATA ON TITLE IX  
COMPLIANCE. YOU CAN  
VIEW IT ONLINE OR ORDER  
REPRINTS OF THE ENTIRE  
SERIES AT [KCSTAR.COM/NCAA](http://KCSTAR.COM/NCAA).  
• THE UNIVERSITY OF IOWA  
CONDUCTED A GENDER  
EQUITY IN SPORTS PROJECT  
ON TITLE IX  
ENFORCEMENT. DATA FOR  
THE PROJECT CAN BE  
FOUND AT [BAILLWICK.LIB.  
UIOWA.EDU/IG](http://BAILLWICK.LIB.<br/>UIOWA.EDU/IG)  
• THE BROWN DAILY HERALD  
CHRONICLED BROWN  
UNIVERSITY'S STRUGGLES  
WITH TITLE IX  
ENFORCEMENT. TO LEARN  
MORE, GO TO  
[NETSPACE.ORG/HERALD/  
LIBRARY/TITLEIX/  
HOMEPAGE.HTML](http://NETSPACE.ORG/HERALD/<br/>LIBRARY/TITLEIX/<br/>HOMEPAGE.HTML)  
CHECK OUT TIPSHEET #854  
FROM THE IRE RESOURCE  
CENTER ON INVESTIGATING  
COLLEGE SPORTS FROM THE  
1998 IRE NATIONAL

• THE WEB SITE ALSO OFFERS FULL TEXT VERSION OF PI'S BIMONTHLY NEWSLETTER, LA RED, AT: [WWW.INVESTIGACION.ORG.MX/LARED/](http://WWW.INVESTIGACION.ORG.MX/LARED/).

• FIND CONTACT INFORMATION ON INVESTIGATIVE REPORTERS FROM LATIN AMERICA AT: [INVESTIGACION.ORG.MX/DIRECTORIO/INDEX.HTML](http://INVESTIGACION.ORG.MX/DIRECTORIO/INDEX.HTML).

# Training difficulties

From page five:

tion, including financial documents and economic data. Many businesses are more cooperative than the government and frequently have information in electronic form.

One advantage for Latin American CAR is that many governmental offices and private firms skipped the era of mainframe computers and work mostly with PCs. Information is in a more accessible and easier-to-use format, as compared to the older and clumsier format found in the United States.

## Success Stories

Many participants in our CAR workshops immediately apply the techniques they learn. A few days after one Access workshop in Brazil, a reporter used the INS database of legal immigrants to the United States from our class to show the "brain drain" of Brazilian executives to the United States. This reporter was no newcomer to CAR, but we have seen many colleagues with no former training work on story ideas or at least dig out those boxes of clippings from under their desk to build a personal database even before the workshop is over.

(This type of enthusiasm can get out of hand. A dear friend of mine has submitted to his spreadsheet analysis everything from how many times his country's lawmakers play hooky to the hobbies, body measures and hair color of the contestants in a beauty pageant.)

Some Latin American colleagues have done very interesting things with their own databases built from paper documents. Reporters from the daily *Público*, in Guadalajara, Mexico, have profiled the arrest and sentencing patterns for drug offenders. The paper got the information from judicial sources though it was like pulling teeth, to later design a database, enter the data and analyze it using Access.

*Público* found that only 5% of the arrests had been the result of police investigations, the rest being the result of random inspections, tip-offs or even traffic accidents in which vehicles carrying drugs were involved. The paper found inconsistencies in sentencing: One a kid caught at home with two dope plants is doing the same jail time as a recurrent smuggler who was captured while transporting dozens of pounds of cocaine.

The Internet also helps reporters reach international sources such as scholars, journalists, and even government sites with information relevant for Latin American countries.

Gerardo Reyes, of the *Miami Herald*, tells a story of one Honduran editor who literally jumped out of his chair when Reyes presented a General Accounting Office Web

## Many media owners are unconvinced about the profitability of CAR.

site. Located there was data on U.S. military aid to the Honduras military. This information is classified in Honduras, he said.

## A few problems

In many newsrooms when owners or managers acquire modern equipment, reporters resist the necessary training either due to fear of technology or because they think their time is too "valuable" to squander in computing lessons. Computers become glorified word processors. Reporters prefer to go to their paper's library or to base their stories on readily available documents and press releases.

At the same time, most media show little interest in paying to have their reporters trained at CAR techniques. Their excuse is the recurring economic crisis, and many of the owners and publishers in the region are still unconvinced of the profitability of modern journalistic resources. One exception, *Folha de Sao Paulo*, sponsors a training program for its staff and for young people fresh out of journalism school and even other fields.

Journalists just getting exposed to CAR, however, often forget about traditional reporting. They fail to turn data into an interesting story. We often read a few dry paragraphs surrounded by tables and graphics or stories that run out of steam three lines after the lead. This will disappear as the novelty of these techniques wears off and especially when data-dazzled editors pick up their pencils and

**Continued on page nine**

From page one:

# Official walls fall

Mexico. Economic reform during the last three administrations has been accompanied by electoral dispute, political assassinations, the appearance of peasant guerrilla movements in the southern part of the country and the rise of left- and right-leaning opposition that threatens the dominance of the governing Revolutionary Institutional Party. Emerging through all of this has been an increas-

---

**We have confirmed that CAR is possible in Latin America and weakens the official monopoly on information. Once the reporting is published, politicians must respond.**

---

ingly participatory citizenry demanding more accountability from politicians.

The media's situation has also changed. Mexico has a large number of dailies — in the capital alone there are more than 20 — but quality journalism has always faced an uphill battle against payoffs and censorship. The same government that could be generous to friendly media could also become a mortal enemy to newspapers running "uncomfortable" stories. But many media are progressively more dependent on public preference and advertisers instead of official subsidies, and government censorship is down.

Our goal at PI is to bring investigative journalism, and in particular CAR, up to a level of sophistication demanded more and more by the public.

When we first offered our workshops on investigative journalism, our first surprise was that despite marked centralism in Mexico, the interest in our project was much greater outside Mexico City. We also started to get requests for workshops from media in other countries in the region, including Colombia, Argentina, Brazil and Peru. PI is now continental, conducting workshops throughout Latin America.

## Interest and Skepticism

Our initial CAR workshops in newsrooms of Mexico and elsewhere in Latin America provoked a mixture of skepticism and interest. Relatively few of our colleagues had used CAR tools such as the Internet, spreadsheets and database managers.

When we showed examples of computer-assisted stories done in the United States, our new friends thought the stories were excellent. But they also brought up all sorts of reasons why this type of journalism couldn't be done in Latin America, from electronic information not being available or reliable to the lack of modern computers in newsrooms.

All agreed that bureaucrats and politicians tend to consider public information as a personal resource, to be shared with the press and society in general only at will. This problem is particularly serious in Mexico, where freedom-of-information legislation is ambiguous and toothless, and in countries where a recent authoritarian past has included stifling journalistic investigation for "national security."

In the past few years, however, we have confirmed that CAR is possible and weakens the official monopoly on information. Reporters can use the Internet to search alternative sources and find the information that bureaucrats and politicians try to hide. They can analyze this information using spreadsheets or database managers and then compare it to official press releases. Once the reporting is published, politicians must respond. Occasionally information is released that was previously controlled.

Some types of official information in Latin America, such as demographic data from the census bureau, health statistics and electoral data, are trustworthy, readily available and helpful for CAR stories. Such information has until now gone unexplored.

We have also used the Web sites maintained by many Latin American governmental offices. These sites offer up much information that would otherwise take days or weeks to retrieve, given the bureaucrats' poor organization or intentional procrastination.

Latin American reporters also make ever-more-frequent use of private-sector informa-

**Continued on page six**

PERIODISTAS DE  
INVESTIGACIÓN AIMS TO  
BECOME A MUST-STOP  
WEB SITE FOR  
JOURNALISTS INTERESTED  
IN INVESTIGATIVE ISSUES  
IN LATIN AMERICA.  
• READ MORE ABOUT THE  
GOALS AND OBJECTIVES OF  
PERIODISTAS DE  
INVESTIGACIÓN AT:  
[INVESTIGACION.ORG.MX/](http://INVESTIGACION.ORG.MX/)  
• PI HAS COLLECTED  
DIGITAL RESOURCES AT:  
[INVESTIGACION.ORG.MX/  
RECURSOS.HTML](http://INVESTIGACION.ORG.MX/RECURSOS.HTML) MOST OF  
THEM ARE IN SPANISH,  
BUT SOME ARE ALSO  
AVAILABLE IN ENGLISH.

# Discovering data

SVITH AND MULVAD ARE REPORTERS WITH MORGENAVISEN JYLLANDS-POSTEN, DENMARK'S LARGEST NEWSPAPER. THEY TEACH CAR AND CO-FOUNDED THE ASSOCIATION OF COMPUTER-ASSISTED REPORTING IN DENMARK AND THE DANISH INSTITUTE FOR COMPUTER-ASSISTED REPORTING. THEY AUTHORED THE FIRST DANISH CAR HANDBOOK, PUBLISHED IN 1998.

• IN APRIL, MORE THAN 60 PARTICIPANTS AND 18 SPEAKERS TOOK PART IN THE GROUP'S SECOND DANISH "CAR-DAY." THEIR ORGANIZATION HAS GROWN FROM 40 TO 140 MEMBERS IN THE PAST YEAR.

• THE GROUP ALSO JUST WON A GRANT FROM THE DANISH PUBLISHER'S ASSOCIATION TO PROCESS DATA AND PUT IT ON THE WEB ([ATWWW.DICAR.DK](http://www.dicar.dk)) AND HELP THE ASSOCIATION WORK WITH BEAT REPORTERS.

**By Flemming Svith and Nils Mulvad**  
*Morgenavisen Jyllands-Posten*

During the last three years, computer-assisted reporting has developed in Denmark from something unknown to an established set of research and analytic skills.

At the outset, there was a defeatist attitude toward CAR in Denmark, but that has changed fundamentally. Active reporters with different areas of expertise have cooperated with each other; a very competent freedom-of-information lawyer has joined the effort; and the management of the Danish School of Journalism has been supportive.

All main Danish media today have heard about CAR and have journalists trained in it. But still only a few newspapers run CAR stories.

## Creating DICAR

In 1997 The Association of Computer-Assisted Reporting in Denmark (FCJ) was founded. Membership has increased steadily. In 1998 FCJ launched the first Danish CAR day.

The association prepared an action program for the establishment of an institute for CAR. The program was to be based on three important principles:

- It would be a nonprofit organization.
- Active journalists would be in control.
- Quick action in the beginning would secure a stable and democratic system for the future.

In 1998 DICAR, the Danish Institute for Computer-Assisted Reporting under the Danish School of Journalism (DJH), was established. Based on experience from Sweden and the United States, the institute filled the need for a more formal and ambitious structure of CAR training and communication.

The statutes and a cooperation agreement with DJH are nearly finished. DICAR has established the first courses and completed a Web page ([www.dicar.dk](http://www.dicar.dk)). The institute has also organized conferences, and it is beginning to receive money. There is so much work to do that it needs to hire a part-time employee. Things are moving ahead.

Two main things are helping the process. One is the experience of American journalists in this field. The other is technological devel-

opment in general, which puts PCs on the desks of journalists and the Internet in a central position for research. CAR is not an obscure realm of nerds but is broadly recognized in mainstream journalism because of the developing computer knowledge in society.

## Pushing for access

One of the main obstacles is access to data. Denmark has 5.2 million inhabitants. Since 1970 the population has had some freedom of information; slowly, a practice of fighting for greater freedom has taken root. Like the authorities of other Scandinavian countries, the Danish government has become largely computerized. But its attitude toward granting access to electronic information has been very restrictive. When information is electronic, there is generally no legal basis for gaining access to it.

Although no right of access to electronic data exists, many authorities are prepared to pass on information, anyway. Moreover, it is possible to invoke other acts and rules in order to generate interpretations that can be used later to formulate precise requests for an alteration of the existing rules.

There have been some very good CAR stories proving that the skills do work in journalism. But still there haven't been sensational CAR stories. No CAR stories until now have been nominated for the Danish version of the Pulitzer Prize, Cavling-prisen.

In the coming year, an important task is making the association and the institute both run stable – on a minor scale much like IRE and NICAR in the United States. Getting the Danish authorities to place an overview of their data on the Internet will be a future task. Other considerable tasks are continuing the dialogue with authorities – both concerning specific experiments with mailing lists and access to electronic documents – and regarding the broader and more fundamental issue of formulating proposals for law reform.

Nils Mulvad can be reached by e-mail at [nmulvad@jp.dk](mailto:nmulvad@jp.dk)

Flemming Svith can be reached by e-mail at [flemming.svith@jp.dk](mailto:flemming.svith@jp.dk)

# CAR on the map

By Dick van Eijk  
NRC Handelsblad

In the past five years, computer-assisted reporting in the Netherlands has developed from an obscure activity of an individual reporter to a formally established practice at two national newspapers. Mapping played an important role in breaking ground for the new professional methods.

To American journalists, it may seem obvious that in their jobs they can use a computer for many tasks other than writing a story. Computer-assisted reporting is something that has existed for thirty years and can be learned at seminars, boot camps and conferences. The situation in the Netherlands, and in the rest of Europe, is entirely different. Only very recently has computer-assisted reporting been gaining momentum in a few countries, notably Sweden and the Netherlands.

I did not know of the American tradition in 1993 when I used a database program (dBase III Plus for MS DOS) to enter the names and positions of more than 2,000 top officials of social-security institutions in the Netherlands. The Dutch parliament was doing a formal investigation on the execution of social-security laws, and we were curious about who actually governed these executing bodies.

From the database analysis, it turned out that there was a relatively small network of unknown officials that met several times a week to determine policy. These officials actually ran the social-security programs. I interviewed most of them and attended, with their permission, a few of their meetings. I wrote two background stories about their hidden world. The stories did not mention that a computer database was used preparing them. Only a few of my colleagues knew.

The database program had proved useful, but at that time I did not realize that the technique had the potential of being an integral part of everyday journalistic practice. That changed during the next year.

## Mapping wish list

To everyone's surprise, local parties gained a strong position in the March 1994 municipal elections. The polls had not predicted the rise of local parties because they concentrated

on national trends and national parties. What we needed, but did not have, was a map showing the percentage of votes collected by local parties in each municipality. Speaking of maps, wouldn't it be nice to show other aspects of election results as well, such as where particular parties were strong or where they had gained or lost the most votes?

That should be possible with a computer, I thought, though I had no idea how to do it. Our national editor encouraged me to find out. I worked largely on evenings and weekends to be able to present such maps on the day after the parliamentary elections, which were only two months away. I made myself familiar with the world of Windows, with Lotus Approach and with the newly discovered mapping program, MapInfo. The application worked, and the day after the elections the newspaper ran a page containing eight maps showing the strength as well as gains and losses of the major parties. The maps were considered a big hit, an experience that would recur in later projects. The explanation, I think, is that the maps showed information not available in such a form before. Their value could be judged without knowing anything about the database analysis behind them.

The power of maps would prove to be an important argument for CAR in the Netherlands. The election maps were discussed on national television, and the professional magazine *De Journalist* wrote about them. We have repeated the map page with every election since and in full color since 1998.

## Checking out America

During the project, I discovered the CAR tradition in the United States. As a reward for my work, my paper sent me to an IRE/NICAR conference in San Jose, Calif., and on a subsequent tour of newspapers with extensive experience in conducting CAR projects. I was ready for a major leap forward.

Our editor was not. It turned out to be difficult to convince others of the potential of CAR. Such projects had to be developed alongside my "real" job of covering the transportation and infrastructure beat. But story showing that the poorest neighborhoods in the country were not situated, as many people

Continued on page nine

TROUW, A SMALLER NATIONAL NEWSPAPER, CAN BE READ ONLINE IN DUTCH AT [TROUW.NL/](http://TROUW.NL/). A LINK THERE TAKES YOU TO NRC HANDELSBLAD AND TWO OTHER DUTCH PAPERS. THIS WINTER TROUW AND NRC HANDELSBLAD TOGETHER ESTABLISHED A FORMAL POSITION FOR A CAR SPECIALIST.



From page seven:

## Locating crime

thought, in the four biggest cities but in mid-size cities in the north and east of the country had a serious impact on urban policy. Teachers at the Utrecht journalism school and a few other journalists started to take notice.

A new editor in the fall of 1996 allowed a little more room to experiment: one day a week. In the meantime, Marjan Agerbeek, an education editor at *Trouw*, a smaller national newspaper, had discovered the power of CAR during a three-month stay at the *Minneapolis Star Tribune* and the *Philadelphia Inquirer*. Together with some colleagues, she conducted some smaller CAR projects, most of which included maps. In the fall of 1997 she did a big one: After a FOIA lawsuit, she acquired the exam results of secondary schools and prepared a supplement on the basis of those results, which until then had not been publicly available data. The project was widely discussed, led to a policy change of the education department and was nominated for the National Newspaper Award.

### CAR advances

*NRC Handelsblad* published a second city neighborhood project, which covered 33 cities and presented full-color maps showing where the "good" and the "bad" neighborhoods were. The maps drew a lot of attention

and were the theme of the first CAR seminar in the country, an afternoon gathering about the use of digital maps and mapping programs. About 40 journalists attended, but so far other newspapers have not picked up the use of these techniques. *Trouw* and *NRC Handelsblad*, however, recognized the potential of CAR and established a formal position for a CAR specialist this winter.

The Utrecht j-school ran a CAR class last year, and other schools are considering doing the same. The emergence of the Internet as a research tool has helped open people's minds to other applications of computers besides word processing. Still, it remains a matter of endlessly repeating arguments and examples. In November 1998 the European Journalism Centre, which provides mid-career training, organized the first spreadsheet seminar for journalists. Seven journalists from three different countries participated.

Of course, compared to the number of journalists attending CAR conferences in the United States, the difference is huge. You have to realize that the pioneers in the United States started in the late 1960s. We started only five years ago.

Dick van Eijk can be reached by e-mail at [eijk@nrc.nl](mailto:eijk@nrc.nl)

From page six:

## CAR wisdom spreads

start doing their job again.

Another problem is that some colleagues don't have a proper knowledge of the field they cover or even basic statistical training. Easy-to-use database and spreadsheet software can lead them to manipulate data incorrectly and reach unwarranted conclusions that might then be printed as news.

We trust that these and other problems will become the exception as Latin American journalism develops more savvy reporters and, of course, readers. Also helpful will be the expanding network of CAR pioneers who, as they refine their methodologies, can share their wisdom with newcomers. This has happened in the United States with the collaboration of IRE and NICAR.

Lastly, it is common in this part of the

world, as elsewhere, for many newly graduated journalists to dream of becoming the latest stars of investigative journalism. And, as elsewhere, vacancies are scarce. But if more young journalists acquire and apply CAR skills to their daily reporting, the quality of Latin American journalism in general will improve. Perhaps, then, we'll finally overcome that old debate on whether investigative journalism is a specialty in itself or if, as Gabriel Garcia Marquez insists, it is a redundant expression.

I would like to thank Alma Delia Fuentes, information and training coordinator at PI, and Lise Olsen who was director from 1996 to 1998.

Pedro can be reached by e-mail at [nanuc@dsi.com.mx](mailto:nanuc@dsi.com.mx)

**CONTACT THE IRE  
RESOURCE CENTER AT  
(573) 882-3364 OR VISIT  
[WWW.IRE.ORG/](http://WWW.IRE.ORG/)  
RESOURCECENTER TO  
SEARCH FOR TIPSHEETS  
COVERING INTERNATIONAL  
REPORTING AND MEXICO  
OR MANY OTHER TOPICS.**

**ALSO SEARCH THE  
LIBRARY OF MORE THAN  
12,000 INVESTIGATIVE  
STORIES ONLINE AT  
[WWW.IRE.ORG/](http://WWW.IRE.ORG/)  
RESOURCECENTER/  
STORYDB.HTML**

**THE RESOURCE CENTER'S  
WEB PAGES INCLUDE  
LINKS TO RESOURCES FOR  
SPECIFIC BEATS AS WELL  
AS INFORMATION ON  
BUYING AUDIO TAPES  
FROM CONFERENCE  
PANELS.**

**KEEP VISITING  
[WWW.IRE.ORG](http://WWW.IRE.ORG) FOR THE  
TIPSHEET OF THE WEEK  
AND THE STORY OF THE  
WEEK.**

# Indexing ease

**MORE SCHOOL STATS:**  
CHECK OUT [NCES.ED.GOV](http://NCES.ED.GOV),  
THE SITE FOR THE  
NATIONAL CENTER FOR  
EDUCATION STATISTICS,  
A FEDERAL AGENCY THE  
COLLECTS STATS ON  
SCHOOLS NATIONWIDE.

**MORE ABOUT THE  
CONSUMER PRICE INDEX:**  
THE BUREAU OF LABOR  
STATISTICS KEEPS  
HISTORICAL AND CURRENT  
DATA ON THE CPI AT  
[STATS.BLS.GOV/CPI/HOME.HTM](http://STATS.BLS.GOV/CPI/HOME.HTM)

**By Neill A. Borowski**  
*The Philadelphia Inquirer*

**Problem:** Your business news department has assembled its annual compilation of executive salaries. Joe Bux at XYZ Corp. makes \$1.4 million a year. Gwen Kash at ABC Inc. makes \$2.3 million a year. And so on. But how do you compare one executive's salary to others?

**Problem:** Once again you're publishing that long and wide list of school test scores. You've heeded the warnings of CAR experts and will not resort to ranking the test scores. How do you give readers a quick look at where their school stands?

**Problem:** The real estate section is ready with its annual look at median home prices by municipality. Ranking the prices would put too much emphasis on the price statistic. But how can readers compare?

**Solution:** In each of these cases, the simple index, a first cousin of percentages, would be the handiest tool. The index is the ideal way to compare all statistics at a glance.

## Comparing Salaries

Consider our executive salary situation above. How can we tell how much one executive is paid relative to another?

An easy approach would be to find the median salary for all of the executives in your survey. The median will be the yardstick with which you'll measure all other salaries. Let's say the median comes out to be \$1.6 million a year. In your index, \$1.6 million = 100. So, what do the individual salaries equal in the index?

Divide the individual salary by the median salary and multiply by 100.

In Joe Bux's case,  $\$1.4/\$1.6 = 0.875 \times 100 = 87.5$  or, rounded, 88.

In Gwen Kash's case,  $\$2.3/\$1.6 = 1.438 \times 100 = 143.8$  or, rounded, 144.

It's a simple calculation in your spreadsheet or database manager. Copy the formula for each of the records, and you have an index.

The "natural sort" for this type of list would be alphabetical by company. This makes the index even more valuable to the reader. You're not ranking by salary. But, at a glance, the reader will see that Kash's salary is far more than the median for the region. In

fact, 44 percent more than the median. Your column label in the newspaper or a footnote would tell the reader that the regional median salary for executives is equal to 100.

You could take it a step further and make the analysis more meaningful by adding another dimension. It always makes sense to ask yourself "Relative to what?" when you do your analysis. In the indexing analysis above, the answer is "Relative to the regional median salary." But this doesn't take into account the size of the company. If XYZ had annual sales of \$500 million a year and ABC had annual sales of \$7 billion a year, the comparison of Bux's and Kash's salaries could be viewed as unfair.

Now take the executive salaries as a percentage of total sales, find the regional median and then index.

In Bux's case,  $1.4/500 = 0.0028$  (in other words, his salary equals 0.28 percent of sales). In Kash's case,  $2.3/7000 = 0.000329$  (or 0.03 percent of sales). These are both tiny numbers, and all of those zeroes after the decimal point would drive the readers crazy.

However, now we find that the regional median ratio of salary to sales is 0.0004. This would be the median after calculating a column of salary-to-sales ratios for all executives. Now we want to derive our index.

In Bux's case,  $0.0028/0.0004 = 7 \times 100 = 700$ .

In Kash's case,  $0.000329/0.0004 = 0.82 \times 100 = 82$ .

The relationship between the Bux and Kash indexes – 700 versus 82 – tells quite a different story than simply indexing their salaries to the median for the region. We've added the element of the ratio to the size of the firm. Of course, these numbers are for illustration purposes. If you make the same calculation with your list of executive salaries, the differences might not be as striking. Or maybe they'll be even more striking.

## Comparing Test Scores

Comparative test scores also could be handled with a simple index. You probably want to publish your list sorted by county first and then alphabetically by the name of the school district. How else might you show how a particular school stacked up against

**Continued on page eleven**

From page ten:

# Weight'em, rank'em

the others without resorting to a misleading ranking?

Most states furnish a statewide average test score. Use this as your base (=100), and derive an index for each of the school test scores. If the state average were 1200 and Sharp School students scored 1275, the index number would be  $1275/1200 = 1.0625 \times 100 = 106.25$ , or 106. At a glance, Sharp parents know their students exceeded the state average. They also can compare Sharp's index number with other schools.

The comparative index still is a ranking. But it is a bit smoother than saying that Sharp School ranked 132nd out of 800 schools.

A way to further smooth the ranking and still maintain some comparative value is to break the schools into quartiles. The quartiles put each school into either the top 25 percent of all schools, the second 25 percent and so on.

Separating the schools into groups of similar schools could provide a more complete picture of how schools compare. One factor to break schools into groups would be income. The only proxy for income most of us have is the reported percentages of students eligible for free or reduced-price lunch. First, break each school into an income subgroup. We have used income quintiles, breaking each school into one of five subgroups based on the percentage/free lunch. Then, within each quintile we break the school scores into quartiles.

This comparison can be an eye-opener. Some wealthier schools that believe they're performing well because they're in the top quartile among all schools discover they're in the bottom quartile when compared to similar schools.

## Comparing Home Prices

In the last situation above, comparing median home sale prices, we simply divide each municipality's median home price into the median home price for the market and multiply by 100 to derive the index number. The market median would equal 100. Other indexes can be derived for other variables in the same table, including the percentage change in the number of homes sold. This is found by taking the municipality's percent-

age change divided by the market percentage change times 100, with the median market percentage change equaling 100.

## Possible Pitfalls: Giving Variables Proper Weighting

Indexes are common in business and economics. The Consumer Price Index as well as the Dow Jones and other stock-market indicators all are indexes.

The simple indexes described above can come in handy. However, caution should be exercised when trying to formulate a complex index. Not every variable should carry an equal weight. And determining the weighting is not an easy task.

Some newspapers have tried to derive their own complex indexes, using several variables. They might want to measure the homeownership cost of a number of municipalities or rate school districts by incorporating several indicators. However, they often simply add up each of the variables to come up with the index number.

The CPI relies on the Consumer Expenditure Survey to select the variable and guide the weighting (for example, what percentage of a "typical" family's budget is devoted to cable television each month). Last year, the Bureau of Labor Statistics went a step further and built a special estimator into the index to help account for substitution effects in consumer budgets, something that the CPI didn't measure well in the past. One substitution example is selecting between ice cream and frozen yogurt based upon which is on sale.

Weighting in other complex indexes is based on factor analysis, a statistical exercise that attempts to determine how a number of factors, or variables, influence an outcome. Factor and regression analyses are used by economists to come up with econometric models, which are forms of complex index.

So, jump with both feet into the world of simple indexing, but be wary of venturing into the complex index without sufficient research.

Neill Borowski can be reached by e-mail address is [nborowski@phillynews.com](mailto:nborowski@phillynews.com).

**SEE RELATED TIPSHEETS FROM 1999 NICAR CONFERENCE ON EDUCATION, REAL ESTATE AND HOME OWNERSHIP STORIES.**

**INVENTIVE WAYS ON ATTACKING EDUCATION STORIES, TIPSHEET #890.**

**HANDLING EDUCATION STATISTICS, TIPSHEET #891.**

**LEARN ABOUT "FLIPPING," A PROPERTY BUYING SCAM, TIPSHEET #880 AND THE APRIL UPLINK.**

**ANALYZING SCHOOL TEST SCORES, TIPSHEET #807.**

**BOROWSKI EXPANDS ON USING CAR IN REAL ESTATE STORIES, TIPSHEET #878.**

**THESE TIPSHEETS ARE AVAILABLE ONLINE AT [WWW.IRE.ORG](http://WWW.IRE.ORG) RESOURCECENTER**

# Fields and forms

By George Landau  
NewsEngin

*Landau's handout was provided at the 1999 NICAR National Conference in Boston.*

Lotus Notes and the Domino server provide one of the most versatile platforms available for designing and deploying database applications that work both from a proprietary client (Notes) and from a Web browser.

---

## The separation of design elements from data makes it easy to modify the design of a database without having to lose or transform any of the data.

---

My favorite guide to Notes development is itself a Notes database - the online documentation that comes with Notes. Because the user help is in a Notes database, you can view it in a number of ways.

You can also create a full-text index to quickly find the entries on any topic. Just open the help menu and choose "Tell me about ... Database Design and Management Concepts" to get started. Here you'll find a detailed presentation of the key concepts, which I'll also introduce here:

**Databases:** Every Notes application uses at least one database, which is a container for actual data along with design elements that control how users see and interact with the data. Databases can also contain agents, which are self-contained little programs that manipulate data in whatever way you can imagine.

**Database design versus documents:** A single database file typically contains all aspects of an application, which means both the design and the actual data. Lotus refers to database records as "documents." Design elements such as views, forms and agents are visible to users with at least "Designer" access rights to the database; everyone else typically can use these elements but can't see how they're put together.

The separation of design elements from data (documents makes it easy to modify the design of a database without having to lose or transform any of the data. And because a master copy of design elements can be kept in an external template, it's easy to experiment with new designs while retaining the ability to revert to prior designs if necessary.

**Forms:** A form in Notes defines the interface through which a user can read and edit individual documents. You can put fields, field labels, descriptive text and pictures on forms. When a user creates or edits a document with a form, the document ends up with that form's fields. You can have any number of forms containing various fields; you might design the application so that certain users see Form A while everyone else sees Form B. A document can have fields that aren't displayed on a form. One of the cool things about Notes is that you can create a new field on a form without having to modify all of your existing data (unlike relational database systems, where you have to alter one or more tables to add a field).

**Views:** Views are a tabular listing of the documents in a database. You can define the "selection formula" for a view to limit which documents it includes. You can insert columns to display the various fields from the underlying data, and you can control sorting of the documents by editing properties of the columns. It might be useful to think of views as predefined queries that show users a particular set of data. Users can further restrict what's listed in views by performing full-text searches.

**Field types:** For the most part, data types in Notes are similar to those in other database software. You can specify text, numeric and date fields. There's also something called a "rich text" field, which can contain an unlimited amount of formatted text, images and file attachments. The only downside to rich text fields is that you can't display them in views.

**Multi-value fields:** With the exception of rich-text fields, any field can be marked by the designer to allow "multi-values." This is like an array or list. You could have a field called "Categories" that contains one or more

**Continued on page thirteen**

NEWSENGIN INC. PROVIDES SOFTWARE DEVELOPMENT AND CONSULTING SERVICES TO THE NEWS INDUSTRY. CHECK OUT OTHER FEATURES ON LANDAU'S WEB SITE AT [NEWSENGIN.COM](http://NEWSENGIN.COM).

- HE INCLUDES A "FREE TOOLS" PAGE THAT HAS A CALCULATOR TO FIND CHANGES IN THE COST OF LIVING BETWEEN YEARS. ALSO, THE SITE HAS A PERCENT CHANGE CALCULATOR.
- YOU CAN ALSO FIND INFORMATION THERE ABOUT PRODUCTS LANDAU OFFERS WITH LOTUS NOTES, INCLUDING SOURCE TRACKER, A CONTACT MANAGEMENT SYSTEM DESIGNED SPECIFICALLY FOR NEWSROOMS. IT SECURELY KEEPS TRACK OF WHAT JOURNALISTS USE: PEOPLE, INTERVIEWS, DOCUMENTS, WEB SITES AND STORIES.

From page two:

## Databases to keep

ownership. Public affairs editor Mike Chalmers said a quick sort of the property database disproved that idea, saving a reporter days of interviews and turning the questions toward wholesalers.

### Long-term CAR:

- Keep databases of property owners and voters. At the *Green Bay Press-Gazette* in Wisconsin, projects reporter Joanne Zipperer said the property database pays big dividends for spot stories and weekenders, yielding owners' names and values of buildings that burn down, people affected by street projects and individual school district issues.

- Think mapping. There's no better way to show editors and readers and to pinpoint the best place to go with eyes and a notebook. Aaron Pilhofer wrote a package last year for Gannett's New Jersey papers on wildly varying property tax rates created by the state's

patchwork of 566 towns and more than 600 school districts. Pilhofer used a database to find high-tax towns next to low-tax towns. He looked at boundary streets and found identical homes in close proximity with wildly different tax bills. He found dozens of examples. "I could have done this by hand, but that would have taken hours or days," he said. "I think it took me less than two hours in all."

- Use existing data. Your newsroom publishes data that would be useful if you could get it into a form that you can count and search. This is the messy world of process reengineering, but it can pay big dividends. Real estate sales, obits, crime incidents and letters to the editor can, in a database, become a great resource for sources, background and trends.

Paul Overberg can be reached by e-mail at [poverberg@usatoday.com](mailto:poverberg@usatoday.com)

### YOUR NEWSROOM

#### PUBLISHES DATA THAT

#### WOULD BE USEFUL IF YOU

#### COULD GET IT INTO A FORM

#### YOU CAN COUNT AND

#### SEARCH:

- REAL ESTATE SALES
- OBITS
- CRIME INCIDENTS

From page twelve:

## Fields and Forms

category names. Multi-values can be a little confusing at first, but they're extremely powerful and easy to use once you get used to them. You can use them to display a single document in more than one place in a view, and you can use them to store data that has

---

**One of the cool things about Notes is that you can create a new field on a form without having to modify all of your existing data.**

---

more than one valid choice. When you write formulas that reference data in multi-value fields, Notes can evaluate the expressions as if the data contained each of the values separately. For example, if a multi-value text field called "jobs" contained "reporter," "photographer" and "editor," any of the following three expressions would be true: `jobs =`

`"reporter" or jobs = "photographer" or jobs = "editor."`

**Agents and actions:** Agents are little programs or code snippets that can act on documents with or without user intervention. Agents can be triggered manually from the Notes menu or from buttons on the "Action Bar" near the top of the screen. In addition to storing code in an agent and running it from an action bar or menu, you can attach bits of code directly to the action. If you plan on reusing the code in more than one part of your application, you're better off writing an agent. Agents can also run automatically on the Domino server triggered by a schedule or by the presence of modified documents in a database. Agents can be scripted using "simple actions" (a limited, wizard-like toolkit for basic tasks like setting field values or sending mail), the Notes formula language (a macro language descended from the 1-2-3 spreadsheet); LotusScript (a dialect of Visual Basic that supports Notes objects) or Java.

George Landau can be reached at (314) 865-4204 or by e-mail at [george@newsenginc.com](mailto:george@newsenginc.com)

LAURA ZUCKERMAN IS A  
NICAR BOOT CAMP  
GRADUATE.  
IRE AND NICAR OFFER  
MANY TRAINING  
POSSIBILITIES, RANGING  
FROM NATIONAL  
CONFERENCES AND  
REGIONAL WORKSHOPS TO  
THE WEEK-LONG BOOT  
CAMPS AND ON-SITE  
NEWSROOM TRAINING.  
VISIT IRE ONLINE AT  
[WWW.IRE.ORG/TRAINING](http://WWW.IRE.ORG/TRAINING) FOR  
A FULL LISTING.

UPCOMING BASIC BOOT  
CAMPS WILL BE HELD IN  
COLUMBIA, MO., ON THE  
FOLLOWING DATES:

- JULY 18-23, 1999
- AUG. 8-13, 1999

CALL (573) 882-2042 FOR  
MORE INFORMATION.

## FIRST VENTURES

# Boating crashes

By Laura Zuckerman

*Daytona Beach News-Journal*

I had downloaded 1997 data from the U.S. Coast Guard and was leisurely winding my way toward a package on recreational-boat crashes when I learned that Florida was within weeks of releasing a report on boat crashes based on 1998 figures.

The law enforcement division within the state's Department of Environmental Protection, in politicized fashion, wanted the release of the data to coincide with the March opening of the 1999 Florida legislature. Knowing that we wanted to go to press with the most current numbers, I decided to sacrifice my ongoing analysis of Coast Guard data in exchange for a daily, 20-25 inch story.

Nice idea, but the state agency wasn't ready for action. DEP officials raised several objections — some more legitimate than others — to my request for electronic information. First, they hadn't finished entering the data. Second, they received regular calls from other newspapers lobbying for the same information. Third, it would steal the thunder from their water-safety presentation to the legislators.

The information was preliminary, so I told the agency that I would emphasize that fact in the story and in any graphics we chose to run.

I was aware we were taking some risks, but, hey, we're journalists, right? We wouldn't shy away from a spot news story, such as covering an accident at the scene, because we didn't have all the information we needed by press time.

## Two timetables

I began negotiating for data the week of Feb. 8. The spokeswoman for DEP tried to persuade me to accept just a few pieces of 1998 data, including the number of boat crashes and boat-related fatalities in our county. Excluded were statewide figures, safety information and crashes by waterway.

I knew we couldn't compare apples to apples nationwide for 1998 but could for past years, based on the Coast Guard data. Nevertheless, I wanted a relatively complete picture of Florida and our coverage area. After several additional discussions with DEP personnel, I faxed them an EFOIA so they

would have in writing the scope of my request and why it was necessary to release the information.

DEP insisted it would take ages for the data-processing center to compile the electronic information. I told them it would take a matter of minutes to copy the database and download it. Clearly, we were both exaggerating, but the point came across: I was impatient to get the figures flowing, and they were just as anxious to slow the process. I didn't want to wait hungrily until the state decided to feed me (and all the other news outlets) information.

Suffice it to say, it was a week of highs and lows, of certain victories and near misses, including negotiating the price from several hundred dollars to less than \$100. By the close of Friday, I had nothing. Finally, after additional legwork by the bureau editor, a fellow reporter and the operations editor, the data was e-mailed Feb. 16.

I imported the data into Access that afternoon and spent most of the night analyzing it and piecing it together with information I'd already gleaned from the Coast Guard. By 1 a.m. Feb. 17, the story was finished.

## Big sport, few referees

The national data already had revealed that Florida was No. 1 for recreational-boat crashes, even though it ranked third in the number of registered boats.

What I learned from the state data was that crashes and fatalities were on the rise, with more in our county from 1997 to 1998 and a slight climb statewide.

More interesting, from a public-safety perspective, was that most people who died in recreational-boat crashes drowned and most of the people who drowned weren't wearing life jackets, which are required in Florida only for some watercraft and some ages. Interviews I had conducted suggested it was politically unpopular to propose legislation that would mandate so-called personal flotation devices, even though marine officers said many drowning deaths could have been prevented.

Legislators I contacted didn't want to talk about life vests, but they wanted to talk about everything else related to recreational-boat crashes. Boating is big in Florida; the laws

**Continued on page fifteen**

From page fourteen:

## Water hazards

that govern the water are weak by degree.

I also broke the data down by waterway, which showed, perhaps not surprisingly, that the Atlantic Ocean and the Gulf of Mexico led the list of boat crashes last year.

We ran the story as a Sunday piece. We had the pleasure of hearing it picked up by other print and broadcast media. I had the disappointment of learning, however, when the state released a report weeks later, that my computer-assisted stories are only as good as my data. The state had mistakenly included several fatalities in the recreational data that should have been dumped into the data on commercial vessels. This caused the number of fatalities we reported for 1998 to be slightly high.

Several times during the boat-crash venture, I experienced computer triumphs, though I had no idea how I had achieved

them. I was equally unsure why other things didn't work with Access, my primary tool, or Excel, a secondary option.

I was interested to read in February a story in Newsweek about the struggle by Microsoft's main "Windows guy," who was videotaped, as part of the government's anti-trust suit, struggling to install programs and access the Internet on an IBM machine.

According to the story by Steven Levy, the Windows guru stumbled, fell, picked himself up, dusted himself off and eventually achieved his goal. I had a similar experience working with the boat-crash data. Computer-assisted reporting is a kind of electronic alchemy with which, through the magic of computers, we try to transmute collections of numbers into relevant, meaningful stories

Laura Zuckerman can be reached by e-mail at [filloislane@yahoo.com](mailto:filloislane@yahoo.com)

---

### FROM THE DATABASE LIBRARY

## Boat Data

By Dawn Fallik  
NICAR Staff

More than 800 people died in more than 8,000 boating accidents in 1997. Many were drunk, most weren't sporting life jackets, and others couldn't swim. Such details about your state's boating accidents are available in the Recreational Boating Database at NICAR for comparison with accidents occurring throughout the nation.

The database dates back to 1969, which means three decades of boating trends. One trend shows that accidents in personal watercraft, such as jet skis, canoes and kayaks, have skyrocketed in the last five years. As a result, some states now require mandatory training.

The database consists of four relational tables:

- **PRIME**, which includes the accidents' basic information (date, time, specific location, cause, the water conditions and whether alcohol was involved) when they involve more than \$500 in damage and/or an injury requiring more than first aid.

- **DEAD**, which includes information on fatalities, such as the victims' age, cause of death, whether they wore a life jacket and

their previous boating experience. This table does not include names and addresses, but fatality reports listing such details are available from state agencies.

- **INJURY**, which includes information similar to that in the "dead" table, only for those injured in an accident.

- **VESSEL**, which includes the type of boat involved in the accident and its operator's experience and blood-alcohol content. NICAR recently won a FOIA battle with the U.S. Coast Guard to have the boat registration and hull identification number included. Now reporters can find boat owners by using the separate boat registration database, also available at NICAR.

The U.S. Coast Guard won't update its national database with 1998 accidents until September. But most state coast-guard agencies have their 1998 databases completed. NICAR can help reporters negotiate through roadblocks thrown up by state agencies, as we did with Boot Camp graduate Laura Zuckerman.

Dawn Fallik can be reached by e-mail at [dawn@nicar.org](mailto:dawn@nicar.org)

**THE NICAR STAFF HAS  
CLEANED UP THE BOATING  
ACCIDENTS DATABASE  
AND PROVIDED NOTES  
POINTING OUT SEVERAL  
DISCREPANCIES,  
INCLUDING SOME  
ACCIDENTS THAT ARE IN  
ONE TABLE BUT DON'T  
LINK TO ANY OTHER  
TABLE.  
DEPENDING ON THE SIZE  
OF YOUR ORGANIZATION,  
NICAR OFFERS BOAT-  
ACCIDENT DATA FOR \$60  
TO \$80 AND THE BOATING  
REGISTRATION DATA FOR  
\$60 TO \$100.**

# Bits, Bytes and Barks

## NICAR data update

Updates of the following data sets are now available through NICAR's Data Library:

- The Occupational Safety and Health Administration's enforcement inspection activity data (1972-1998) details inspection reports on accidents, hazardous substance injuries, workplace violations and fines.

- The 1998 Alcohol Tobacco and Firearms' Firearms Licensee database provides information on licensed gun and explosives dealers.

- The Center for Disease Control's AIDS Surveillance database (1982-1997) contains cases reported to state and local health departments.

To order, call NICAR at (573) 884-7332.

## Campaign Finance Workshop

Staff members from IRE's Campaign Finance Information Center will be training journalists on techniques for investigating political contributions at a workshop July 23-25, 1999 in Chicago. For updates on the workshop see [www.campaignfinance.org/training.html](http://www.campaignfinance.org/training.html).

## IRE Regional Conference

West-coast subscribers may be interested in the IRE Regional Conference coming up in Los Angeles on September 9-12, 1999. For updates on that conference and information on other IRE and NICAR conferences, see [www.ire.org/resources/conferences.2](http://www.ire.org/resources/conferences.2)

## Boston Tipsheets

For those who could not make it to the National Computer-Assisted Reporting Conference held in Boston this past March, the IRE Resource Center has archived all 90-plus tipsheets from the seminars.

Search our database at [www.ire.org/resourcecenter](http://www.ire.org/resourcecenter) to see what's offered and to order copies of the tipsheets. You can get a complete set for \$175 (that's \$285 for nonmembers).

## Upcoming Boot Camps July 18-23 and Aug. 8-13

CAR Boot Camps are weeklong, on-campus seminars held at the Missouri School of Journalism in Columbia about five times per year. They start on a Sunday afternoon and continue to Friday afternoon.

Boot Camps give journalists a jumpstart in computer-assisted reporting techniques. They 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 more information, call IRE and NICAR at (573) 882-2042 or visit the IRE Web site at [www.ire.org](http://www.ire.org).

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

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