December 1998

Uplink

A newsletter of the National Institute for Computer-Assisted Reporting

UNDERMINING DEFINING

Suspect semantics

By Craig R. McCoy and Tom Torok
The Philadelphia Inquirer

Philadelphia sure is a law-abiding place.

It's a city where people aren't held up at gunpoint. Instead, they have "disturbances," or, at worst, are subject to "threats." Victims aren't beaten or stabbed. They are deemed mere "hospital cases," coding for the police taking someone to the hospital. Pocketbooks aren't ripped off; they're booked as "missing property." Thieves don't steal cars. Instead, the car's forgetful owners lose them, and the helpful cops write up the missing vehicles as a "try and locate."

With artful classifications such as these, Philadelphia's police department for years has been fudging statistics for the FBI's influential Uniform Crime Reporting (UCR) program. This, in turn, has permitted mayor after mayor to boast of the city's low crime rate.

Until the whole statistical deck of cards came undone. In a series of articles over the last year, *The Philadelphia Inquirer* has challenged the city's crime statistics, using computerized reporting and a lot of old-fashioned, shoe-leather interviewing to document flaws in the numbers.

The finagling has had a big impact, both for the citizenry as a whole and, on a very personal level, for victims:

- At this point, no one really knows what the true crime picture is for the city. In stages, police over the last year have disowned the citywide crime stats for 1996, 1997 (twice withdrawn) and the first half of 1998.
- The scandal has damaged Philadelphia's attempt to follow New York's lead and switch over to police deployment based on computerized trend-spotting and mapping.
- Some "downgradings" – the term for converting a major crime to a minor one – have cut off cases from detectives, leaving suspects free to strike again.
- Victims have lost their claim for compensation under a state program that helps pay for uninsured medical bills. No crime, no victim, no compensation.

The stage for our reporting was set when Philadelphia Mayor Edward Rendell, finally responding to requests from the *Inquirer* and the *Philadelphia Daily News*, released seven years of underlying crime-by-crime data. This massive computer file provided details on more than 700,000 crimes deemed serious by police from 1991 to 1997. Philadelphia is one of few major American cities that release

Continued on page four

CORRECTIONAL CARE

Prisoner health

By David Heath
St. Louis Post-Dispatch

While in an Alabama prison, 18-year-old Calvin Moore melted before the eyes of correctional doctors and nurses. Moore, a burglar who suffered from severe mental illness, lost 56 pounds in less than a month – more than one-third of his body weight. On Feb. 21, 1996, he died from dehydration and starvation in the care of St. Louis-based Correctional Medical Services, the industry giant in the rapidly growing field of privatized correctional health care.

CMS said it provided Moore with "appropriate and compassionate care." A state report attributed the death to natural causes. However, a renowned expert on forensic medicine and pathology from the University of Colorado described Moore's death as "a homicide resulting from criminal negligence."

Word of Moore's death prompted the *St. Louis Post-Dispatch* to spend five months looking at the privatized correc-

Continued on page eight

Inside Uplink

For the patrollers of computer-assisted reporting, the crime beat is rampant with opportunity. This issue of *Uplink* samples ways in which CAR lends force to crime investigations – from detailed analysis of the FBI to city police departments to negligence and abuse behind bars.

Craig McCoy and Tom Torok of *The Philadelphia Inquirer* document how local police have "downgraded" crime statistics to alter favorably the city's crime rates.

David Heath of the *St. Louis Post-Dispatch* uncovers national examples of negligence by private health care providers for prisoners.

Jim Hopkins of *The Courier-Journal* (Louisville, Ky.), after lengthy data entry, proves a pattern of inmate abuse followed by weak disciplinary action against the guards who inflicted it.

David Burnham and Susan Long of TRAC share lessons – applicable to stories on all levels of the criminal justice system – from their in-depth portrayal of the FBI.

Uplink

December 1998
Volume 10, Number 11
A newsletter of the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston
brant@nicar.org

MANAGING EDITOR
Brent Johnson
bjohnson@nicar.org

SENIOR CONTRIBUTING EDITORS
Len Bruzzese
len@ire.org
Sarah Cohen
sarah@nicar.org
Richard Mullins
richard@nicar.org

COPY EDITOR
Jeanine M. Davis

ART DIRECTOR
Wendy Charron

STAFF
Britton Clapp
Dawn Fallik
Jason Grotto
Ben Lesser
Ted Peterson
Noemi Ramirez
Aaron Springer
Neil Warner
Alistair White

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
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.

Investigating the FBI

By David Burnham and Susan Long
TRAC

The FBI is an extraordinarily secretive organization that routinely denies Congress, reporters and the public comprehensive information about the matters it investigates, the number of individuals who are indicted and convicted as a result, or the distribution of its staff in different parts of the country.

For many decades, this absence of concrete information has made it nearly impossible to judge the bureau's actual performance. Given America's faith in the theory of accountability, the public acceptance of the FBI position is surprising. It's as if General Motors could sell stock without publishing an annual report certified by an independent accounting team.

Fortunately, however, the FBI's steadfast resistance to public scrutiny can now be challenged. This is possible because of the Freedom of Information Act, the availability of computerized data analysis programs and the recognition that very detailed information about the enforcement activities of the bureau are recorded by the Justice Department's Executive Office of United States Attorneys (EOUSA). Previously, our knowledge of FBI enforcement practices was almost entirely limited to the announcements and leaks from the FBI and the DOJ about cases they wanted to publicize.

Stepping stones

The process of conducting an investigative audit of the FBI – or for that matter a local school district, sheriff's office, or state police – almost always involves the same basic steps, which often overlap as an investigation proceeds.

First, obtain comprehensive data about the individual enforcement actions of the agency you have targeted for examination – the records of its inspections, investigations, arrests, referrals, audits, prosecutions, and so on.

Second, check the accuracy by comparing the data obtained from the targeted agency with parallel information compiled by a separate institution. Confidence in the general accuracy of enforcement information increases when two or more independent agencies produce similar counts.

Third, sum the case-by-case information so that significant aspects of the agency's work product can be quantified. How many drug prosecutions were there? How many civil rights prosecutions? How many convictions? How many were sentenced to prison? An examination of this summary information often reveals unusual patterns and trends. Has there been a sharp and unexplained surge, or decline, in one particular kind of enforcement? Do records indicate the agency is more successful in processing one group of suspects than another? Is the agency focusing a large proportion of its efforts on an unexpected subject? Do staffing patterns appear reasonable?

Because we began by collecting referral-by-referral data about all kinds of enforcement actions, however, we could examine a highly focused investigative subject.

Fourth, on the basis of these preliminary revelations, interview responsible administrators and other knowledgeable persons to gain their insights and criticisms. When the data revelations are challenged, the challengers should always be asked to produce their data.

FOIAing data's turnkeys

We have followed these four steps for more than thirty years, initially working alone from the somewhat different perspectives of investigative reporter and research scholar. Since 1989, when we joined forces as co-directors of the Transactional Records Access Clearinghouse (TRAC), we have sought to make this kind of information widely available. (See *http://trac.syr.edu*)

A basic source of TRAC's enforcement information is the EOUSA. Through the aggressive use of the FOIA, TRAC has ob-

Continued on page three

From page four:

Toward accountability

tained referral-by-referral information about all federal criminal matters and what happened to them from the period when this system of records was developed in the mid-1970s through fiscal year 1997. TRAC also has obtained data from such agencies as the IRS, ATF and Office of Personnel Management, as well as the federal court system and the GAO.

In late 1995, after obtaining comprehensive data from the Justice Department about the criminal enforcement activities of the entire federal government, we decided the time had come to systematically examine the FBI, which in many ways is the most powerful and least accountable agency in the United States.

From the beginning, we viewed our probe of the FBI as a two-tracked project. One track resulted in the publication of an investigative article that comprised most of the August 11/18, 1997 issue of *The Nation*. (See *http://www.thenation.com/issue/970811/0811burn.htm*) The second track was more unusual: the organized data that served as the backbone of our investigation became a Web site where anyone could independently examine what the bureau is doing, where it is doing it and how the agency's current activities compare with those of past. The site has more than 30,000 pages of maps, graphs, tables and other material.

Door-breaking over deskwork

From our examination of the Justice Department's enforcement data, GAO studies, court records and interviews with working and retired FBI officials, federal prosecutors, judges, criminal defense lawyers and other experts, we slowly developed a wide range of findings about the FBI. One key conclusion: the FBI spends a substantial proportion of its investigative resources on matters that often could be handled by local police and relatively little on problems that can be dealt with best at the national level and by the federal government. In 1996, for example, bureau investigations resulted in 11,865 convictions. Half of them involved drugs, bank robbery and fraud against banks, frequently by credit card. By contrast, according to department data, only 148 of the convictions, 1.2 percent of the total, involved medical fraud.

Interviews suggested that agents much preferred the high-adrenalin, kick-in-the-door cases over the demanding and sometimes tedious work required to convict white-collar criminals. The data proving the FBI has paid relatively little attention to medical fraud – a national problem of immediate concern to millions of elderly Americans – is genuine news by itself, confirmation of the rule that what an agency does not do often is more important than what it does. The health fraud data were also valuable because they measured uniquely the broad dimensions of a significant failure of government. Because we began by collecting referral-by-referral data about all kinds of enforcement actions, however, we could examine a highly focused investigative subject.

From 1992 to 1996, the Justice Department reported that federal prosecutors disposed of 142 FBI matters it had classified as involving internal security and terrorism – cases of highest concern by definition. But the data showed only 22 percent of these 142 matters resulted in a conviction. In other words, more than three-quarters were declined by the prosecutors, dismissed by the judges, or ended in a not-guilty verdict. Of those matters that the Justice Department decided to decline, about two-thirds were tossed out because the prosecutors concluded the evidence was too weak or legally flawed.

Marshaling the marshals

Though difficult because enforcement agencies have different responsibilities, reporters can also marshal the Justice Department to examine the FBI's effectiveness. One approach is to compare how the FBI and another agency handle a specific category of cases. Both the FBI and DEA, for example, enforce the nation's drug laws.

Through the sophisticated use of the Justice Department data now available through TRAC, backed up by careful reporting of the old-fashioned kind, the nation's most important law enforcement agency can now be subjected to the kind of examination it has long required and almost never received.

David Burnham and Susan Long can be reached at (315) 443-3563 or (202) 544-8722

"THE FBI: A SPECIAL REPORT," WHICH APPEARED IN *THE NATION*, WAS A FINALIST IN THE MAGAZINE CATEGORY OF THE IRE AWARDS.

THE ADDRESS FOR TRAC'S WEB SITE IS HTTP://TRAC.SYR.EDU. THERE IS NO CHARGE FOR ACCESS TO TRAC'S SITES ON THE FBI, IRS AND DEA. DEPENDING ON THE SERVICES SELECTED, THERE ARE A RANGE OF CHARGES FOR TRACFED, A FOURTH SITE WHERE REPORTERS CAN GET INFORMATION ABOUT ALL CRIMINAL ENFORCEMENT ACTIONS OF THE FEDERAL GOVERNMENT FOR THE LAST FIVE YEARS. THE DATA ARE AVAILABLE FOR THE COUNTRY AS A WHOLE, INDIVIDUAL DISTRICTS, JUSTICE DEPARTMENT PROGRAM CATEGORIES (E.G., PORNOGRAPHY, MEDICAL FRAUD, DOMESTIC TERRORISM), BY LEAD CHARGE, OR BY AGENCY. TRACFED HAS TWO SERVICES – EXPRESS AND ANALYZER, WITH WHICH REPORTERS CAN CREATE DATA SLICES FOR DETAILED ANALYSIS.

McCoy writes:

Our reporting has also triggered follow-up coverage from other media. CNN, NBC's Dateline, NPR, the new 60 Minutes II, the All Points Bulletin crime-news Web site (http://www.apbonline.com/911/1998/11/16/phillystats1116_1.html), and a fledgling RealAudio-based Internet news service called WorldStream Communications, among others, are all taking a look at the issue.

The 1991-97 crime database can be reached at http://home.phillynews.com/crime. The site also incudes a discussion of the data, a collection of many of the news stories discussed in this article and other crime-related databases for our region.

From page one:

Downgraded crimes

such detailed information, besides totals made public twice a year by the UCR program.

Using Access and Excel, an *Inquirer* team began to slice and dice the data. In some ways, our focus was on the dogs that weren't barking – crimes that weren't being booked.

Boom and bust thefts

Among our early articles was a straight forward account, pegged to the FBI's national release, of the city's overall crime rate for 1997. Even as we wrote it, we noticed a weird trend in the data. When crimes were analyzed by month, Philadelphia had a puzzling drop in car thefts in the first half of the year, followed by a boom in the second. The change was without precedent for Philadelphia; we could not explain it as seasonal.

We then took our analysis to the city's police chief, John F. Timoney, for comment.

The numbers at issue had already been withdrawn from the FBI count once. In late 1997 the *Daily News* first reported that the city had been counting crimes based on when they entered data into their computer system rather than when crimes occurred. This method made comparisons between years and among cities haphazard at best.

Timoney, who inherited the mess when he took command of the force in March 1998, had released new 1997 figures after tallying them again for a regular Jan. 1-Dec. 31 calendar year.

Once he had digested the *Inquirer*'s new analysis, Timoney announced in May 1998 yet another inquiry into the 1997 figures, less than a week after their re-release. In September, he provided the results: for the first half of the year, police had by mistake dropped 3,000 property crimes – mostly stolen cars – from the city's count.

Novelties and oddities

As we continued to explore the database, other puzzling patterns emerged. A key finding was Philadelphia's remarkably low count of aggravated assaults compared to murders. While almost every major city routinely books 50 such serious assaults for every murder, Philadelphia was logging just 15 per murder. Criminologists pointed us to data showing that murder and serious assaults rose and fell together in that ratio in most cities – a not surprising tidal effect if you think of a murder as an assault that goes too far.

Another oddity: attempted burglaries had just about vanished in Philadelphia. For example, the data let us see that two police districts, with a total population of 138,000 people, reported a total of only *six* attempted break-ins in 1997. Shoplifting figures had been similarly low-balled. So had thefts of auto licenses, a particularly troublesome Philadelphia problem.

Major UCR scandals have flared repeatedly. Experts have been warning that the problem is growing more virulent as departments put more stress on the numbers, both to fight crime and to rank commanders.

Though Philadelphia's pervasive culture of numbers chicanery may seem unusual, there is no doubt that the problem prevails in municipalities large and small. Since the system's founding nearly 70 years ago, major UCR scandals have flared repeatedly. Experts have been warning that the problem is growing more virulent as departments put more stress on the numbers, both to fight crime and to rank commanders. In the last year statistical controversies have been sparked in New York, New Orleans, Baltimore, Atlanta and Boca Raton.

Shoe-leather series

As we worked over the data, we also interviewed serving and retired police from beat cops to commanders; victims and victims' advocates; town watch and civic group leaders; and academic and policing experts nationwide. We also consulted court records and ordered scores of individual incident reports.

The department's long-time culture of "going down with crime" – as the cops called it internally – began to surface. Detectives

Continued on page five

From page six:

Accuracy efforts

and street cops told us how crimes were blown off to lighten their workload. Others spoke of constant pressure from the top, of district commanders and even more senior officials seeking to raise performance ratings by putting a thumb on the scale of justice.

Our work culminated Nov. 1-2 in a series that was actually sparing in its use of numbers and charts. Instead, we told story after story of actual victims and how the crime against them had been downgraded. One article explored violent crime; another tackled crimes against property. Among the more dramatic evidence were graphic reproductions of complete police reports that had been rewritten to downgrade the crimes. In the rewrites, key information vanished, including descriptions of suspects and weapons.

Wrongdoing and response

The issue has engendered considerable outrage in Philadelphia. For the first time, the city controller, the top financial watchdog for the city, is planning an audit of crime figures. U.S. Attorney General Janet Reno has asked her aides to examine the Philadelphia figures.

Timoney has sworn the 1998 figures will be accurate, "if it kills me." He has transferred two captains whose numbers were questioned and created an aggressive auditing unit to check figures. So far their investigation suggests that as many as 5,500 crimes were buried in the first half of 1998 – one out of every 10.

After our series, Timoney ratcheted the pressure higher. He announced that undercover police would pose as crime victims to "sting" officers mishandling paperwork.

Character overload

While examining the Philadelphia figures, we also looked at national data. We drew, in part, upon unpublished FBI statistics that the agency provides on magnetic tape (available from NICAR) that includes in its entirety crime data collected for the UCR program – much more detailed information than what is available in print in the well-known *Crime in the United States* publications. One problem is the FBI sends the data in text that runs 8,000 characters across, four times the length that most databases can handle. Using a Visual Basic script, Torok helped break down the data into usable sections. (*Note from the Database Library: NICAR processes the raw FBI data so software using the .dbf format can handle them.*)

At the paper's request, a sociologist at Temple University had built an ArcView GIS map of Philadelphia, which tagged each of the city's 400 police sectors by neighborhood. With this in hand, for one article graphics artist Matthew Ericson allocated crime by sector to each of Philadelphia's 56 neighborhoods. We ranked them by violent crime and property crime rates.

Online database

With leads flowing in, we are still at work on the issue. In one recent development, we posted on *The Inquirer*'s Web site, Philadelphia Online, the 700,000-crime database that underpinned the entire effort. It took considerable grooming to get the database into usable form. Using Visual Basic, Torok "parsed" the chaotic address field into discrete block and street fields. His program correctly separated all but 15,000 of the crimes. Ericson then wrote a program in Perl (CQ) that solved another particularly bad 6,000 addresses or so. We also deployed scripts that fixed more than 100,000 misspellings of street names.

Using Microsoft SQL Server, Torok then created active server pages on the Web site that permit readers to search the database rapidly for crimes on their block over seven years, on any one day, and by other methods. His program ingeniously separates and indexes the data by search method to produce quick results from an immense dataset. It received more than 24,000 page views its first day online. For comparison, a well-browsed *Inquirer* sports story receives around 3,000 "hits."

One of our first responses came in an e-mail, edited slightly for confidentiality: "I looked for an assault," a browser wrote us. "It must have been downgraded since it is not listed at all. It infuriates me that an assault is not reported. The two men who did this meant to cause bodily damage, but their act of cowardice won't be known because the police department is playing games with our lives."

We're looking into it.

Craig McCoy can be reached by e-mail at mccoyc@philly.infi.net

Tom Torok can be reached by e-mail at Tom.Torok@phillynews.com

OTHER EXAMPLES OF WEB SITES DEVOTED TO CRIME STATISTICS AND MAPPING INCLUDE:

- **THE CHARLOTTE OBSERVER AT HTTP://WWW.CHARLOTTE.COM/CRIME**
- **THE EVANSVILLE COURIER AT HTPP://COURIER.EVANSVILLE.NET/CRIME**
- **OMAHA WORLD-HERALD AT HTTP://WWW.OMAHA.COM/OWH/CRIMEREPORT**

MCCOY ALSO SUGGESTS VSITING THE WEB SITE OF THE SAN DIEGO POLICE DEPARTMENT AT HTTP://WWW.SANNET.GOV/POLICE/CRIME-FACTS/INDEX.HTML

COMING IN THE JANUARY ISSUE OF UPLINK ARE TWO MORE STORIES ON USING COMPUTER-ASSISTED REPORTING FOR COVERING CRIME:

- CAROL NAPOLITANO OF THE OMAHA WORLD HERALD WILL EXPLAIN THE APPLICATIONS OF HER PAPER'S SPECIAL REPORT ON CRIME, WHICH WAS HEAVY ON BOTH FEATURES AND STATISTICAL MAPS
- NAOMI AOKI OF THE VIRGINIAN-PILOT WILL DISCUSS HOW TO USE MAPPING FOR CRIME STORIES

"WE DON'T FILE OUT TAPES"

On the data offensive

By Bob Warner
Philadelphia Daily News

Police departments around the country have been happy to take public bows for declining crime rates. But in many cases, they're still refusing to provide news organizations with electronic data that would allow closer scrutiny of their records.

As of early 1998, police departments in the nation's three biggest cities – New York, Los Angeles and Chicago – were unwilling to release any data to reporters describing individual crimes. Houston's police department said it would be willing to release electronic data, but as of last January nobody had asked. Of the five biggest cities, only the Philadelphia police were providing the media with electronic data on local crime – and only after a two-year fight requiring intervention by a progressive mayor.

The Philadelphia Daily News began asking the police department for various pieces of electronic data – drug arrests, stolen cars, etc. – in 1994. With the exception of a small database describing city homicides, our requests were turned down, flatly and repeatedly, by the department's chief counsel and its day-to-day spokeswoman. In early 1996, we began asking for more comprehensive data – the past five years of criminal offenses reported to police, known to a small group inside the department as the "offense tapes." Once again, we were turned down. "We don't give out tapes," the department's spokeswoman told us.

Pennsylvania's public records law offered us little support. Written in 1957, it does not mention computers or electronic data. It contains a troublesome exemption for any records disclosing the institution or progress of a criminal investigation. Since the police claimed to investigate all Part 1 crimes reported to it, its lawyers would occasionally argue that this exemption covered any records on crimes reported to police.

Media allies

But we had an ally in the city's mayor, Ed Rendell, whose top aides had the common sense to manage the news media by providing information, not withholding it. Unofficially, they had long since agreed to the principle that documents considered public records on paper should also be considered public records when they're kept in an electronic format. We asked the mayor's office to review our dispute with the police.

It was a ticklish matter for the mayor's office, which had repeatedly been accused by the police officers' union of trying to run the department from the second floor of City Hall. But some sort of message was actually delivered to Police Commissioner Richard Neal, telling him he'd have to give us some data or provide reasonable explanations for withholding them.

Called and cornered

We still had trouble getting Neal to return our calls. Finally, we cornered him as he was kicking off a pedestrian safety campaign. That led to a series of three meetings, going through the department's record layouts and negotiating precisely what data would be turned over. The newspaper agreed to live without victim's names and addresses, but the department agreed to identify the blocks or intersections where crimes had occurred.

The agreement nearly fell apart with Neal's next phone call. Our data was ready, he said, but the *Daily News* would have to pay $2,089 – in advance – to cover the department's programming costs. Reluctantly, we agreed.

It didn't take long to find evidence of the department's bogus crime-counting practices: thousands of crimes from 1995 had been included in the data used to come up with 1996 crime totals for the FBI. It soon became clear that the department's annual crime counts were based on when they fed data into their computer system rather than when the crimes occurred.

Our stories forced the department to withdraw crime figures previously submitted to the FBI for 1996 and 1997. More importantly, they became Exhibit A in a hot public debate over the effectiveness of the police department and the competence of Commissioner Neal. Five months later, Neal submitted his resignation. His successor, John Timoney, made clear on taking office that reform of the department's criminal data system would be one of his top priorities. Timoney has followed through.

Bob Warner can be reached at (215) 854-5885.

CRIME WITHOUT PUNISHMENT

Inmate abuse

By Jim Hopkins
The Courier-Journal

Shortly after midnight on Jan. 7, 1998, a 34-year-old inmate died after a struggle with guards in the Jefferson County jail in Louisville, Ky. Within days, the county coroner ruled his death a homicide, setting off multiple investigations by the jail, the police, prosecuting attorney and *The Courier-Journal.*

Early on, we reported that three of the five guards involved had prior on-the-job disciplinary records, including one who had been accused of inmate abuse. That piqued my curiosity and a question that formed the basis of our inquiry: How many other officers in the department had been disciplined, and for what violations?

The answer came the following month when we reported that nearly 90 percent of the jail's 350 guards had been disciplined for violating workplace rules. One-fifth had been found guilty of major offenses such as inmate abuse, insubordination and sleeping on the job. Serious punishment, such as termination, was rare. One officer had been disciplined 48 times – more than any other – for violations that included pummeling an inmate in the face and then trying to cover up the incident. Management had tried to fire him twice but failed after he appealed to an employee grievance board.

Open-records accommodations

To reach our conclusion, we relied heavily on computer analysis and Kentucky's open-records laws to review all the disciplinary records of current and former jail-house staff. With our findings in hand, we turned to the Internet to find jail-management experts to place the situation in a national context.

In the wake of our report and a follow-up piece that looked exclusively at inmate abuse, jail managers negotiated a new contract with the union representing guards that is supposed to make it easier to punish officers who don't follow the rules while keeping an eye on the jail's approximately 1,600 inmates.

When we first asked the jail for all employee disciplinary records, we were presented with a loose-leaf log that listed every employee offense, the date, a brief description of the offense, the recommended punishment, and the outcome. The log became our road map.

Unfortunately, the log wasn't computerized. Using Paradox, I entered each incident

I included every scrap of information in the log because I wasn't sure what would be important until I was finished.

and eventually created a 2,236-record database. I included every scrap of information in the log because I wasn't sure what would be important until I was finished.

I detected patterns almost immediately: Some employees had been punished dozens of times; management seldom recommended termination; and employees often got their punishments overturned on appeal. For example, our analysis showed that nearly 34 percent of all recommended suspensions had been reduced by the employee grievance board. About 70 percent of the recommended firings taken to the board were reduced to lesser punishments.

In the case of the inmate who had died, three of the five officers involved had prior infractions that could have led to termination. But in only one instance was firing recommended, and that recommendation was tossed out during the grievance process.

The employee grievance panel, it turned out, gave management and the Teamsters union an equal number of votes when deciding cases appealed to the board. Our analysis showed that virtually all discipline that involved a suspension or termination were appealed to the board. Anxious to avoid a prolonged appeal if they stuck to their guns, managers repeatedly struck bargains with the union.

Data entry demands

The lack of a pre-existing computer data-

Continued on page fifteen

NEW! EDITOR BOOT CAMP

THESE NEW COMPUTER-ASSISTED REPORTING BOOT CAMPS FOR EDITORS ARE INTENSIVE THREE-DAY WORKSHOPS TAILORED TO THE NEEDS OF NEWSROOM MANAGERS. THIS INCLUDES TOP EDITORS OR NEWS DIRECTORS, MANAGING EDITORS, AMES, ASSIGNMENT EDITORS AND OTHER EDITORS DIRECTING REPORTERS. THESE BOOT CAMPS WILL TEACH EDITORS THE THINGS THEY NEED TO KNOW TO MAKE CAR SUCCESSFUL IN THEIR NEWSROOMS. THEY'LL EXPERIENCE JUST ENOUGH HANDS-ON WORK TO UNDERSTAND WHAT THEIR REPORTERS ARE TACKLING -- AND WHAT MORE IS POSSIBLE. THE FIRST OF THESE BOOT CAMPS IS SCHEDULED FOR FEB. 5-7, 1999. REGISTRATION INFORMATION WILL BE AVAILABLE ONLINE SOON AT *WWW.IRE.ORG*

THE POST-DISPATCH'S FINDINGS WERE PRESENTED SEPT. 27 IN A 12-PAGE SPECIAL SECTION, TITLED "HEALTH CARE BEHIND BARS: DEATH, NEGLECT AND THE BOTTOM LINE." THE SERIES IS STILL AVAILABLE ON OUR WEB SITE AT *HTTP://SPECIAL.POSTNET.COM/PRISONCARE/INDEX.HTML*.

IRE FEED 4 – CALL FOR TAPES

FOR INFORMATION ON SUBMITTING A TAPE OF YOUR BROADCAST INVESTIGATION FOR THE NEXT COMPILATION OF INNOVATIVE STORIES, WHICH WILL INCLUDE COMPUTER-ASSISTED RESEARCH AND OTHER STATE-OF-THE-ART APPROACHES, CONTACT MARK LAGERKVIST AT (516) 393-1299 OR STUART WATSON AT (910) 310-2025.

From page one:

Disciplinary data

tional health-care industry. Reporters found more than 20 cases in which inmates died as a result of alleged negligence, indifference, under-staffing, inadequate training or cost-cutting by for-profit HMOs.

Information obstacles

The search for information was not easy. State correctional agencies were generally uncooperative, concerning the backgrounds of CMS physicians. We knew from anecdotal evidence that some had checkered pasts. The problem was obtaining a list of all CMS doctors and matching them against a database of disciplinary actions by state medical boards.

Although the data we needed existed, it required sweat to obtain and make usable. A database of unprofessional and incompetent behavior by medical doctors, called the National Practitioner Data Bank (*www.npdb.com*), is maintained by the U.S. Department of Health and Human Services. However, federal law prohibits the names of the doctors in the database from being divulged to the public. That information is only disclosed to hospitals, health care entities, state licensing boards, and, in some cases, attorneys.

The Missouri State Board of Registration for the Healing Art is more open with disciplinary data. The board even posts the physician license database, which includes a field of disciplinary actions (*www.ecodev.state.mo.us/pr/*).

Index end-run

The only comprehensive database available to the public is provided by Public Citizen, a consumer group founded by Ralph Nader. The data come in a four-volume set of books priced at $307.50. It can be ordered from the Web site (*www.publiccitizen.org*). Regional listings are available for $23.50.

The books provide an abstract of actions taken against doctors (including arrests), actions taken by state licensing boards, and actions taken by the Drug Enforcement Administration. While it is a gold mine of information, it is organized by state rather than by doctors' names – a problem since many doctors with disciplinary histories move around. Even the index was difficult to navigate, because it was grouped by type of action.

We solved these problems by scanning the 75-page index and writing a short program in Visual Basic to parse the fields into a database. We posted a search form on a password-protected Web site that we shared, at first, only with members of the project team. When the series ran, we made that database available to everyone in the newsroom on our Intranet. The Web site allows reporters to search doctors by name.

Prying with contracts

Our bigger problem was obtaining the names of doctors working for CMS in order to match them against our data. Our computer-assisted-reporting specialist, Julie Lucas, contacted 26 states in which CMS does business to ask for a list of doctors working there and the contract the state signed with CMS.

In virtually every case, the states at first refused to provide us with the list of doctors, claiming that they did not have that information or directing us to contact CMS directly. (CMS declined to share the data as well.) We did routinely obtain the CMS contracts, however. In some cases, those contracts provided the leverage we needed to pry the list of CMS doctors from the states. In nearly every contract, CMS was bound to provide the state correctional department with a current list of their doctors. We used that contract language to persuade several states that the doctors' names were a public record.

Even so, two states, Virginia and Arkansas, claimed that we weren't entitled to the information because their open record acts apply only to state residents. In actuality, Virginia's law specifically states that newspapers circulating in the state are entitled to public records. We proved that we sell some papers there. In Arkansas, we enlisted the help of Jeff Porter, the computer-assisted reporting editor at the *Arkansas Democrat-Gazette*, who managed to get the list of doctors within a day.

Our efforts paid off. We had numerous hits on our database from the list of doctors we received from the states. In fact, doctors with questionable backgrounds became a major theme in the series.

David Heath can be reached by e-mail at dheath@gw.pulitzer.net

Death on the job

Bill Muller
The Arizona Republic

Bill Muller attended the NICAR basic boot camp in August 1997.

It was one of those moments that every reporter dreads. I woke up in the middle of the night staring at the ceiling. After analyzing an OSHA database, I had just finished writing a story about workplace deaths in Arizona. In cases where there was a serious safety violation, I had identified 108 deaths since 1992 and established the median fine per fatality: about $3,000.

But now I tossed and turned. What if there were more than 108 deaths? What if the data was flawed? Our analysis would be worthless.

Luckily, though I already had written the story, its run date was still weeks away.

The next day, I called the director of the state industrial commission, who I had come to know well during the project, and told him of my misgivings. Without hesitation, he agreed to have a member of his staff go through all workplace deaths since 1992 and make sure we had them all.

We didn't.

Undocumented deaths

Due to some reporting problems on OSHA's part, 26 deaths were not in our analysis. Using data provided by the commission, we added the deaths into the Excel spreadsheet that we had used to calculate the median. A funny thing happened: The median was still about $3,000.

In picking median for analysis, we had done ourselves a favor. When we added the missing cases, some came in above and some came in below, and the median stayed about the same.

Still a good lesson. Some other findings:

• From 1992 until mid-1997, 79 of the 134 worker deaths investigated by OSHA in Arizona led to citations for serious safety violations. Just 12 fines were more than $10,000.

• Though OSHA fines related to workplace deaths might start high, the appeal system is designed to reduce them. In fact, an analysis of initial fines and current fines showed that they dropped by about 40 percent, saving the companies involved about half a million dollars.

• For widows and widowers of workers who die on the job, Arizona ranks low for survivor benefits. State law protects companies from wrongful death lawsuits, saying essentially that employees covered by workers' compensation cannot recover damages in the event of injury or death.

Though OSHA fines related to workplace deaths might start high, the appeal system is designed to reduce them.

Early experiments

Naturally, getting the answers wasn't as easy as we first thought. It was my first time dealing with a fairly large database – about 30 MB, with four tables and 43,000 records in the largest table. The data weren't laid out quite the way I thought they would be.

For example, when I first calculated the fines per death (sorted in Access, then crunched in Excel), I noticed that each record included only one company name. This puzzled me because I knew several cases in which more than one company had been fined in connection with a death. To get the death cases, you had to link the accident table with the OSHA table by case ID number. Each time I asked for cases with '1' in the degree field (the death code), I got the same list and only one company name for each deceased worker.

I decided to experiment. I sorted the data by date and then by address, looking for investigations in the OSHA table that occurred on the same date and at the same address as deaths in the accident table. Essentially this worked. To make sure the cases were related (and not just OSHA inspectors spotting unrelated violations on the job site), we checked the related activity numbers that appear on the case records available on the OSHA Web site. We also had the industrial commission check our work. They found

Continued on page fifteen

CONFERENCE AUDIO TAPES:

AUDIO TAPES FROM THE MARCH 1998 SESSIONS AT INDIANA CAR ARE NOW AVAILABLE FOR ORDERING FROM SOUND IMAGES, WHICH CAN BE REACHED AT (303) 649-1811. FOR MORE INFORMATION OR A DOWNLOADABLE ORDER FORM, POINT YOUR BROWSER TO *WWW.IRE.ORG/RESOURCES/NICAR/CONFERENCES/INDIANA/AUDIO.HTML*

AUDIO TAPES FROM THE JUNE 1998 IRE NATIONAL CONFERENCE IN NEW ORLEANS ARE ALSO AVAILABLE. FOR MORE INFORMATION, POINT YOUR BROWSER TO *WWW.IRE.ORG/RESOURCES/CONFERENCES/NEWORLEANS/AUDIO.HTML*

BROADCAST AWARD WINNERS ON VIDEO:

THE VIDEO COMPILATION OF THE 1997 IRE AWARD BROADCAST WINNERS IS AVAILABLE FOR SALE. THE 2 1/2-HOUR VHS TAPE COSTS $10. CALL THE RESOURCE CENTER AT (573) 882-3364.

IF YOU HAVE A PARTICULAR PROBLEM YOU WOULD LIKE TO SEE ADDRESSED IN A FUTURE TECH TIP, SEND AN E-MAIL TO BRENT JOHNSON AT BJOHNSON@NICAR.ORG

UPDATED DATA:

THE DATABASE LIBRARY HAS UPDATED THE FOLLOWING DATA SETS:

- CONSOLIDATED FEDERAL FUNDS REPORT, 1986-1997 – FEDERAL EXPENDITURES OR OBLIGATIONS IN STATE, COUNTY, AND SUBCOUNTY AREAS, INCLUDING U.S. OUTLYING AREAS AND WASHINGTON, D.C.
- MEDICAL DEVICE REPORTS, 1984-1997 – INFORMATION ON DEFECTIVE OR POTENTIALLY DEFECTIVE MEDICAL DEVICES ALREADY ON THE MARKET
- FEDERAL RAILROAD ADMINISTRATION'S ACCIDENTS DATABASE, 1991-1997 – INFORMATION ON U.S. RAILROAD ACCIDENTS, INCLUDING WEATHER CONDITIONS, TYPE OF ACCIDENT, NUMBER OF CARS, AND THE LIKE

Programmers aplenty

By Steve Doig
Arizona State University

A big part of the fun for most CAR folks is the challenge of solving puzzles. Our projects tend to be one-of-a-kind things, so we're constantly facing new data formats with strange layouts (at least to us), undocumented codes and other idiosyncrasies. Even though no one else in the newsroom understands how difficult the struggle was, we get a rush of satisfaction from finally beating an unruly dataset into submission.

When tackling such data puzzles, a great resource is the NICAR-L list. Even if you've never faced variable-length records or packed hexadecimals or YYYYDDMM dates or bitmaps or other weirdness, chances are good that others have. And perhaps the most instructive thing about the flurry of postings often generated by an interesting question is the wide range of clever solutions offered. Over and over, we discover that – in CAR – there are a lot of ways to get from here to there.

A case in point was the "ugly data" contest I tossed onto NICAR-L this summer. It not only provided a stimulating puzzle for some of our best propeller-heads, but the results also proved the versatility of the CAR toolbox.

I found this puzzle in a big tape dataset sent to me by a reporter who needed it in a more useful format and on CD-ROM. The data, which came as a courts database involving civil suits, offered several kinds of ugliness:

- The structure was hierarchical: For each case there was one case record and then as many individual party records as there were parties to the case. But the case record didn't say how many party records would follow. Nor was there a common field, like case number, to link case and parties.
- Each "line" of the data was composed of exactly 100 of the above case and party records, with no delimiters between records. Thus, a "line" might be anywhere from 4,500 to about 6,000 characters long. If the last case still had more parties when the 100-record limit was reached, the rest would spill across to a new "line."
- The file layout we were given was wrong. It didn't mention that each case record began with "0026" and each party record with "0055." Later, I figured out that this undocumented feature was supposed to be the length of each record type, though why the file creators felt the need to repeat this for every record is beyond me.
- Nor were the supposed record lengths of 26 or 55 characters reliable. Some, but not all, records ended with one, two or three carets, for no apparent reason.

It took me an afternoon of noodling with SAS, my database program of choice, to figure out a way to read the data and transform it into something useful. I've been on a long campaign to convince CAR power users that SAS is better than more commonly used programs such as Access or FoxPro, particularly for ugly data like this. So I created the "Ugly Data Contest," urging anyone interested to tackle it with whatever tools they wanted. I was sure their struggles, and the relative simplicity of my SAS program, would win some converts.

Over and over, we discover that – in CAR – there are a lot of ways to get from here to there.

The roads to solution

I won't go into all the details of the results here (you can check the NICAR-L archives of August 4-10, 1998, by going online at *http://www.nicar.org/nicarl.html*). But I was surprised and delighted – despite my pro-SAS agenda – at the wide variety and inventiveness of the solutions that came in.

Along with my SAS program, there were FoxPro programs from Greg Reeves of the *Kansas City Star* and Griff Palmer of the *Daily Oklahoman*; a Visual Basic solution from David Heath of the *St. Louis Post-Dispatch*; a stand-alone Visual Basic GUI from Dan Keating of *The Miami Herald*; a Word-and-Excel method-in-concept (that I called the "ugly solution" winner!) from Russ Clemings of the *Fresno Bee*; a Visual FoxPro/SQL approach from Jeff Porter of the *Arkansas Democrat-Gazette*; and a text editor/Monarch solution from Dick van Eijk of NRC Handelsblad in Rotterdam, The Netherlands.

Continued on page fourteen

ON THE INTERNET

Crime data online

By Heather Browne
NICAR

Covering crime is not for the faint of heart. Crime statistics, for their part, can be daunting or difficult to locate. The vast array of criminal statistics out there can be hard to navigate, but the Web helps narrow it into categories. Some sites offer answers to frequently asked questions for data that needs explanation or provide tutorials for analyzing the data. The following are just a few of the many sites that offer everything from interesting crime tidbits to vast databases. They are good places to start and can lead to more specific sites.

The FBI Home Page
www.fbi.gov

This site is the premier source for the latest in crime stats. Not only can you check out who's who on the "Most Wanted" list, but you can also get the most recent version of "Crime in the U.S.," which summarizes Uniform Crime Reports (UCR) from agencies throughout the country. (Click on "Crime Statistics") It includes trends and thorough explanations, which save time. Also available is the National Incident-Based Reporting System, the latest wave in reporting crime.

The Bureau of Justice Statistics
www.ojp.usdoj.gov/bjs

The U.S. Department of Justice offers much in the way of statistics on both victims and offenders. Findings from the National Crime Victimization Survey can flesh out FBI data – the survey may account for unreported crimes. This site also offers spreadsheets in trend data, information on courts and sentencing, and much more. It incorporates different data sources, so it is very useful in tracking crimes through the various stages of criminal justice. A lot of the data is available to order.

The Justice Information Center
www.ncjrs.org

This site offers in-depth information that supplements FBI and other crime statistics. It includes statistics in areas such as homicide reports, arrests, and juvenile court, or victim information. For specific topics, such as "Crime Prevention" or "Juvenile Justice," it lists links to relevant Web sites, mailing lists, and articles. Definitely check out its extensive list of sites for criminal justice statistics at *http://www.ncjrs.org/statwww.htm.*

The National Consortium for Justice Information and Statistics
www.search.org

If the data you're looking at appears to be a scrambled mess, visit this site for background. It offers information on courts, NIBRS, and can lend technical support. The consortium site gives dates for conferences and symposia in crime statistics and criminal justice. It can also link you to the Clearinghouse, another wealthy source of information for reading and analyzing data on law and order in the U.S.

FedStats
www.fedstats.gov

This site is simply awesome. It links you to any government agency and is simple to follow. It calls itself "one-stop shopping" for statistics, programs, and policy. Maintained by the Federal Interagency Council on Statistical Policy, federal statistics from the FBI, DEA, and many more agencies are just a click away. FedStats also lists agency contacts.

Organized Crime
www.crime.org

Not the Mafia, though you could probably link to them. This site, which is easy to move around in, provides tutorials on reading crime statistics and links to all the major sites. It helps answer questions on crime data and points out possible inaccuracies. Check it out for nothing else but the always enlightening "criminal quote." The designer, Regina Schekall, plans to develop the site further.

Number Crunchers Precision Journalism
www.crunch.com

This site is devoted to journalists. It "crunches" the numbers, based on the adage that "journalists don't do math." It's got some "top secret" FBI files, now available under the Freedom of Information Act. You can look up the rap sheet of Al Capone, for example. A glitzy, colorful site, it offers data and spreadsheets on crime trends.

National Archive of Criminal Justice Data
www.icpsr.umich.edu/NACJD

This site offers a wide array of background information, technical support, and data. If you need to wade through all that, go ahead, but the site offers a link straight to the downloadable data, found in outline form. Just click on the subject you'd like to see: corrections, police, victimization surveys, and more.

Continued on page fourteen

THE WEB SITE FOR THE CAMPAIGN FINANCE INFORMATION CENTER (WWW.CAMPAIGNFINANCE.ORG) IS A VALUABLE RESOURCE FOR ANYONE INVESTIGATING CAMPAIGN FINANCES. IT INCLUDES:

- **DOWNLOADABLE DATABASES OF CONTRIBUTION DATA AND LINKS TO SEARCH ENGINES MAINTAINED BY NON-PROFIT ORGANIZATIONS AND STATE BOARDS OF ELECTION**
- **STORIES FROM PRIOR ISSUES OF *TRACKER***
- **STORIES EXPLORING CAMPAIGN FINANCE AT THE LOCAL, STATE AND FEDERAL LEVELS**
- **INFORMATION ON THE CFIC-L MAILING LIST**
- **TIPSHEETS ON COVERING CAMPAIGN FINANCE**
- **LINKS TO POTENTIAL SOURCES FOR CAMPAIGN FINANCE STORIES AND OTHER RELEVANT SITES**

NICAR's Database Library can convert data from all types of government systems. The staff can take nine-track tapes, 4mm DAT tapes, 3480/3490E cartridges, print image files or TR1/TR2 tapes, and put them onto CD-ROM in the database format of your choice. NICAR Staff cleans and converts the data, and performs integrity checks. Prices are based on the size of your news organization. Call the database library at (573) 882-0684 for more information.

POWER BOOST:

The Database Library recently purchased a 3480/3490/3490E tape drive. In the library's experience processing local and federal data sets, it appears that many agencies are moving toward the 3490E format, which is written at 36-track. If you receive data in this format, NICAR can now process it.

Calculating inflation

By Sarah Cohen
NICAR

Almost every comparison using dollars over a number of years should be adjusted for inflation. It's simply not fair to talk about salary or budget increases since the 1970s without building in inflation.

Not enough reporters do this, and that's a shame. It's also now inexcusable. The data is sitting on the Web, as are tools that make the simple calculation for you.

A recent question from an IRE member prompted me to test a variety of these inflation-adjustment (or cost-of-living) calculators on the Web. I was looking for results consistent with my own calculations, documentation that let me verify how they'd come to an answer, and a simple interface that clearly guided me to enter data for the answer I expected.

Not many of the tools met my challenge.

And the winner is...

Only George Landau's NewsEngin tool passed in all categories: accuracy, documentation and ease of use. Landau, a longtime IRE member, also confirms the date of the last update, letting you know his data is up-to-the-minute accurate. The instructions of some calculators were so indecipherable that you could easily make a mistake filling them out. Others introduced rounding errors or slightly erroneous data into their calculations.

Any site that doesn't keep its data up to date and doesn't provide details on how it performs its calculations ought to be eliminated from your list of bookmarks.

The rounding error isn't very worrisome. In fact, if rounding makes such a big difference in an answer, there's a good reason to report the answers at a general level. There's no reason to imply a false sense of precision when doing work based on indexes with bigger flaws than a calculation would introduce.

But any site that doesn't keep its data up to date and doesn't provide details on how it performs its calculations ought to be eliminated from your list of bookmarks. It's just not

Comparison of Web calculators

Here's a comparison of the Web sites I tested:

Name & Link	Accuracy	Documentation	Interface	Comments
AIER Cost-of-Living Calculator: www.aier.org/colcalc.html	**	***	***	Good interface that clearly identifies what you're doing. Answers are slightly off, probably due to an odd version of the CPI used.
George Landau's inflation calculator: www.newsengin.com/neFreeTools.nsf/CPIcalc	***	***	***	Longtime IRE member Landau works from up-to-date versions of BLS databases. Good interface lets you adjust forward or backward.
NASA's Java version of an inflation calculator: www.jsc.nasa.gov/bu2/inflate.html	**	*	**	Allows adjustment using GDP deflator as well as CPI. Good links to other calculators. Allows downloadable spreadsheet of indexes. Appears to round interim answers.
Robert Sahr's inflation factors: www.orst.edu/Dept/pol_sci/fac/sahr/cv1998.gif	*	***	*	Explanation of how to use an index, but you do the calculation based on his estimates of inflation rather than up-to-date official figures.
S. Morgan Friedman's Inflation Calculator: www.westegg.com/inflation/	**	*	*	Difficult interface with no documentation.
The Dismal Scientist: www.dismal.com/toolbox/cpi_index.stm	**	**	*	Good explanations of how the CPI is compiled but no documentation of how it is applied. Answers are within rounding error, but it's unclear why they're slightly off.

From page twelve:

Web as backup

worth the effort to worry that you've filled it out incorrectly, that it's using the correct formula, or that the data is out of date.

How the calculators calculate

Generally, you type in a value like a salary. You tell the inflation adjuster what year the salary applies to and in what year you want it expressed. Then you receive an inflation-adjusted figure.

Almost all of the calculators assume you want to inflate previous year dollars. So if you type in $25,000, 1975 and 1997, you'll get what a salary of $25,000 in 1975 would be today, after inflation.

That's usually how we do it, even though the instructions on official Web sites imply you ought to use some other base year. The reason is that readers, viewers and listeners can picture what a dollar is worth today. Few accurately remember what it was worth a long time ago.

On your own

Instead of depending on someone else, most reporters who deal with dollar values like budgets, salaries or home values ought to learn how to compute inflation-adjusted figures themselves.

Then you can leave the Web tools where they belong – as a check on your own work.

It's not hard. I promise. Here's the formula:

(CPI Now / CPI Then) * Value then = Old dollars inflated to compare in new dollars.

To get the CPI, just visit the BLS home page at *http://www.bls.gov www.bls.gov*. Then choose "Data," and "Most requested series." Almost everyone uses the CPI-U, or the national CPI for all urban consumers, as their deflator. It doesn't matter which base period you use; they all use the same underlying data.

Instead of depending on someone else, most reporters who deal with dollar values like budgets, salaries or home values ought to learn how to compute inflation-adjusted figures themselves. Then you can leave the Web tools where they belong – as a check on your own work.

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

Spreadsheet

Here's what a spreadsheet of formulas and answers would look like:

	A	B	C	D	E	F	G
3	Year	Federal budget	Pct change	CPI	Infl adj formula	Infl adj result	Pct change aft adj
4	1988	$1,064,489	NA	118.3	=(d$13/d4)*b4	$1,444,214	NA
5	1989	$1,143,671	7%	124.0	=(d$13/d5)*b5	$1,480,316	2%
6	1990	$1,253,163	10%	130.7	=(d$13/d6)*b6	$1,538,888	4%
7	1991	$1,324,400	6%	136.2	=(d$13/d7)*b7	$1,560,692	1%
8	1992	$1,381,681	4%	140.3	=(d$13/d8)*b8	$1,580,612	1%
9	1993	$1,409,414	2%	144.5	=(d$13/d9)*b9	$1,565,474	-1%
10	1994	$1,461,731	4%	148.2	=(d$13/d10)*b10	$1,583,049	1%
11	1995	$1,515,729	4%	152.4	=(d$13/d11)*b11	$1,596,289	1%
12	1996	$1,560,512	3%	156.9	=(d$13/d12)*b12	$1,596,317	0%
13	1997	$1,601,235	3%	160.5	=(d$13/d13)*b13	$1,601,235	0%
14							
15	Total		50%				11%
16							

UPCOMING TRAINING:

MAY 10-13, 1999 – IRE AND NICAR AND THE UNIVERSITY OF NORTH CAROLINA SCHOOL OF JOURNALISM WILL HOST AN ADVANCED SEMINAR ON USING STATISTICS IN COMPUTER-ASSISTED REPORTING AT THE UNIVERSITY OF NORTH CAROLINA.

INSTRUCTORS WILL INCLUDE PHILIP MEYER, UNC PROFESSOR AND AUTHOR OF "THE NEW PRECISION JOURNALISM," AND SARAH COHEN, IRE AND NICAR TRAINING DIRECTOR.

AN OPTIONAL SESSION ON THE BASICS OF THE STATISTICAL SOFTWARE SPSS WILL TAKE PLACE SUNDAY, MAY 9 (MOTHER'S DAY).

MAY 16-21, 1999 – BASIC BOOT CAMP IN COLUMBIA, MO.

CALL (573) 882-2042 FOR MORE INFORMATION OR VISIT THE IRE WEB SITE AT *WWW.IRE.ORG/RESOURCES/CONFERENCES/TRAINING/*

TO STEER YOUR CAR PROJECTS IN THE RIGHT DIRECTION, ORDER "COMPUTER-ASSISTED REPORTING: A PRACTICAL GUIDE" BY BRANT HOUSTON. IT PROVIDES A GATEWAY TO THE FIELD FOR BOTH BEGINNING AND EXPERIENCED JOURNALISTS WHO NEED TO LEARN ESSENTIAL SKILLS AND GET UP TO SPEED IN THE COMPUTER AGE. A PC DISK WITH DATA EXERCISES IS INCLUDED.

ORDER THE BOOK FROM NICAR AND INVESTIGATIVE REPORTERS AND EDITORS FOR $19.98 (EARLIER PRICED AT $26 FOR IRE MEMBERS OR $30 FOR NON-MEMBERS). CALL (573) 882-2042. FOR A LIST OF OTHER TITLES, POINT YOUR BROWSER TO THE IRE BOOKSTORE AT *WWW.IRE.ORG/PUBLICATIONS/BOOKSTORE/*

SUBSCRIBE ONLINE:

YOU CAN NOW SUBSCRIBE TO *UPLINK* ONLINE. POINT YOUR BROWSER TO *WWW.IRE.ORG/RESOURCES/NICAR/UPLINK*

From page eleven:

Abundant statistics

Or search by letter of the alphabet. There are also forums to which you can post questions.

Security on Campus
www.soconline.org/STATS/index.html

Statistics concerning crime on college campuses is hard to come by, but this site makes it easy. It gives background information and offers links to UCR data and various university crime tables.

Heather Browne can be reached by phone at (573) 884-7711

From the mailbox

On the NICAR-L mailing list, David Milliron of *The Atlanta Journal-Constitution* recently suggested these Web sites for the crime beat:

National Clearinghouse for Criminal Justice Information Systems
http://www.ch.search.org/

Crime Data Sources in Criminal Justice
http://www.ncwc.edu/~toconnor/data.htm

Uniform Crime Reports: County Level Data
http://fisher.lib.virginia.edu/crime/
1995 appears to be the latest year.

Sourcebook of Criminal Justice Statistics
http://www.albany.edu/sourcebook/
Data about all aspects of criminal justice in the U.S. Includes more than 600 tables in Adobe Acrobat .pdf format.

Justice Research and Statistics Association
Statistical Analysis Centers with Web Servers
http://www.jrsainfo.org/sac/sacsites.html

Finding forensic information on the Internet
http://www.bart.nl/~geradts/lit.html

Crime and Justice Electronic Data Abstracts
http://www.ojp.usdoj.gov/bjs/dtdata.htm
The Department of Justice Bureau of Justice Statistics recently updated this massive database to include selected data through 1996. Lotus spreadsheets (. wk1) are available in 55 subjects covering crimes and arrests, criminal justice, and general demography.

UN Crime & Justice Information Network
Statistics and Research Sources
http://www.ifs.univie.ac.at/~uncjin/stats.html

Crime Mapping Research Center
http://www.ojp.usdoj.gov/cmrc/
Wealth of data on crime mapping and agencies that have put their crime maps on the Internet.

Students' Report of School Crime: 1989 and 1995
http://www.ojp.usdoj.gov/bjs/abstract/srsc.htm
"It compares findings from the 1989 and 1995 SCS on student reports of victimization, drug availability, street gang presence, and gun presence at school."

Law Enforcement Sites on the Web
http://www.ih2000.net/ira/ira2.htm
Possibly the largest collection of law enforcement information on the Web.

Federal Bureau of Prisons
http://www.bop.gov/
Includes directory of facilities, Fact Sheets with summary statistical data on federal prison populations by gender, race, ethnicity, age and other demographic characteristics.

David Milliron can be reached by e-mail at david-milliron@david-milliron.net

From page ten:

Seeking SAS converts

But my favorite solution used an entirely unexpected tool: PERL, the "print extraction and report language" loved by Unix wizards and cgi-scripters. Matt Ericson of the *Philadelphia Inquirer's* web operation did it with these five lines:

• Put a new line after any party record

```
perl -i -pe 's/(\d\d\d\d\d\d)(\d\d\d\d)([CP])/\1\n\3/g' ug.txt
```

• Put a new line after any case record

```
perl -i -pe 's/(\d\d\d\d\d\d\w\w\w\w\w\w)(\d\d\d\d)([CP])/\1\n\3/g' ug.txt
```

• Fix any line that still starts with a 4 digit number, which seems to appear after the ^s

```
perl -i -pe 's/^\d\d\d\d//;' ug.txt
```

• Kill any ^s

```
perl -i -pe 's/\^//g;' ug.txt
```

• Merge the cases and parties

```
perl -i -pe 'if (/^C/) { $c=$_; $_=""; chop($c) } else { $_="$c  $_" }' ug.txt
```

I still say SAS is best for handling the widest range of ugly data and other CAR problems. (I am, however, starting to learn some PERL!) But the contest showed it's possible to make almost any CAR tool do almost any CAR task. As the old saying goes: "When all you have is a hammer, every problem looks like a nail."

Steve Doig can be reached by e-mail at steve.doig@asu.edu

From page seven:

Grievous reductions

base was the biggest problem; virtually all the jail records were paper. I don't enjoy data-entry work, but there didn't seem to be any other way. It took about three solid days, but in cases like this I do the work myself because I can see trends early on, which helps me draft the story as I go. Inputting data myself also makes me feel more confident about accuracy.

The employee log was only a starting point; it included virtually no details from individual disciplinary cases. To learn more about inmate abuse and excessive force, for example, I had to isolate the names of all 34 officers accused of such violations in the past. Then I requested their complete personnel files. That led to a second story, which concluded that although inmate deaths at the hands of guards were rare, incidents of excessive force and abuse were not.

Lessons learned

What's true in traditional reporting also is true in computer-assisted reporting: A newspaper's greatest ally is the open-records law. In this case, however, the records told only part of the story. Although jail managers knew they were having a hard time enforcing the rules, they didn't know how serious the problem was because they had never taken a close look at their own records. That's where computer analysis proved invaluable.

Finally, database stories can be pretty dull stuff without concrete examples. Part of what made our two stories compelling was our focus on case studies, including the officer who had been disciplined 48 times. Using the database as a guide to inmate abuse cases, for example, I described in detail situations where guards had beaten inmates held down by other officers, falsified reports to cover up their actions, and gotten off scot-free.

Jim Hopkins can be reached by e-mail at hopkinsj@courier-journal.com

From page nine:

Swim in the OSHA

one related case that didn't turn up, mostly because it was added to the case file a month after the accident.

Scanning for the striking

After generating our initial list of 108 deaths, we searched the data for cases that would be interesting. We looked for small fines, large fines, and large fines that hadn't been paid (by checking the "penalty remit" field in the data). I also went through the newspaper library, scanning for details on various cases.

We found striking cases, including a small company that had dodged a $47,000 fine by declaring bankruptcy and another that was fined just $7,500 after a worker had been decapitated. To find the relatives of these victims, many of whom had scattered, we used AutoTrack, locating one relative in Rochester, N.Y., and another by sending letters to her last two known addresses.

Also, after an editor asked the question, I went back to the data to find out if we could document the total amount that OSHA fines had been lowered during the appeal process. Fortunately, the data allowed such a calculation by including "current penalty" and "initial penalty" fields. Because of ongoing appeals, the fines associated with more recent cases tended to change, so we updated them using the OSHA Web site establishment search. (See *http://www.osha.gov/cgi-bin/est/est1*)

But 26 fatalities were still missing from the records. The reason, as far as we can tell, was that state officials had not filled out the accident summary when logging the case. So while the accident appeared in the main OSHA table, it was not in the accident table, which contained the degree field and the name of the victim.

I used the records provided by state OSHA to update my chart, and the rest is history.

By the way, you might want to replace that frayed electric cord.

Bill Muller can be reached by e-mail at mullyb@aol.com

TO SEARCH THE COLLECTION OF CONFERENCE HANDOUTS AVAILABLE FROM THE IRE RESOURCE CENTER, DIRECT YOUR WEB BROWSER TO WWW.IRE.ORG/RESOURCES/CENTER/HANDSEARCH.HTML

TO SEARCH THE LIBRARY OF MORE THAN 12,000 INVESTIGATIVE STORIES AVAILABLE, GO TO WWW.IRE.ORG/RESOURCES/CENTER/SEARCH.HTML

THE RESOURCE CENTER CAN BE REACHED AT (573) 882-3364 OR BY E-MAIL AT RESCNTR@IRE.ORG

THE ARIZONA REPUBLIC'S OSHA PROJECT, "DEATH ON THE JOB" CAN BE ACCESSED ONLINE AT HTTP://WWW.AZCENTRAL.COM/NEWS/OSHA/INDEX.HTML

Bits, Bytes and Barks

IRE and NICAR on the crime beat

As you can see from this issue, reporting on crime in the United States is becoming more sophisticated each year as reporters try to give it context and depth. IRE and NICAR are working with several groups to improve reporting on crime.

One project involves the Berkeley Media Studies Group and the Missouri School of Journalism. The project is trying to show journalists how to strengthen daily reporting on crime and violence by making use of datasets and historical information. It also tries to link the reporters to public health experts who study the topic intensely.

Lori Dorfman from Berkeley and veteran reporter Jane Stevens are leading the effort. Both have some radical ideas about daily reporting on crime and violence and are taking their message into newsrooms.

The role of IRE and NICAR is to help provide meaningful hands-on training and examples of stories that use combinations of datasets and shoe-leather reporting. IRE and NICAR also hope to work with a newly created association on crime reporting to provide manuals, tipsheets and booklets about covering different aspects of the beat: the cop shop, the court, the prison.

If you have ideas or want to help on either of these projects, please contact Brant Houston by e-mail at brant@ire.org

Conferences abound

Information on upcoming IRE and NICAR national and regional conferences is now available on the IRE Web site. Visit the IRE Web site at *www.ire.org/resources/conferences/*

The **National Computer-Assisted Reporting Conference**, presented by IRE and NICAR along with *The Boston Globe*, will be held March 11-14, 1999 in Boston. You may register online or download an early registration form at *www.ire.org/resources/conferences/boston*

The schedule for the conference is being finalized. We hope to offer new classes in SAS basics and data management and PERL scripting. We plan to reinstate basic Fox training using Visual Fox and continue popular basic classes, a special session for broadcasters, and Intranet training.

Contact Sarah Cohen, IRE and NICAR training director, with suggestions for panel topics and hands-on classes, plus ideas for who might speak or teach. Cohen can be reached by e-mail at sarah@nicar.org

The **Investigative Reporters and Editors National Conference**, co-sponsored by *The Kansas City Star* and KCTV, will be held June 3-6, 1999 in Kansas City. You may also register online or download an early registration form at *www.ire.org/resources/conferences/kansascity*

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

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