December 2000

Uplink

RACIAL PROFILING

The race factor

By Dan Browning
Minneapolis Star-Tribune

Minneapolis police say they don't use racial profiling to stop alleged lawbreakers. They rely on experience, aided by computer-driven analysis that points them to trouble spots. But their computer-assisted hunches lead them to more black suspects proportionately than whites.

Are police using good judgment? Or are they trapping relatively more black criminals because they're sweeping up higher percentages of blacks than whites?

Philip Meyer, a journalism professor at the University of North Carolina at Chapel Hill and one of the leading proponents in the use of quantitative analysis by journalists, suggested a kind of cost-benefit study.

Assuming that the judicial system is fair, even if police rely heavily on race or ethnicity when deciding whom to arrest or ticket, and the proportion of minorities who are convicted of crimes is similar to the proportion of whites who are convicted, then the police strategy might be reasonable, Meyer said.

Conversely, if blacks are arrested or ticketed at higher rates than whites for similar crimes, but are less likely than whites to be convicted, then police judgment is arguably unsound. Such a finding would support claims that racial prejudice was resulting in unfair, inefficient and costly enforcement practices, he said.

Testing the theory

The Star-Tribune tested this theory and on July 23 published a special report, "Presumed Guilty Until Proved Innocent."

When studying racial profiling, or police bias, reporters ideally would study stops. But Minnesota lacks such data. Next might come searches, or tickets. But that data, too, was unavailable. That leaves bookings.

Hennepin County is the one Minnesota jurisdiction using a computer system that allows easy tracking of an offender's case from arrest to adjudication. It was suitable for our analysis because nearly two-thirds of all blacks in Minnesota live in that county. (Almost half of all blacks in the state live in Minneapolis.)

We took the county's jail data from 1994 through 1999 and linked it to sentencing data, then flagged each booking record that resulted in a sentence of some kind.

The court data covered sentences only; it did not include diversion, probation before

Continued on page twelve

HEALTH CARE

Nursing mistakes

By Mike Berens
Chicago Tribune

MAUDE was my first source.

She led me to a largely hidden world of death and injury inside America's hospitals. She provided scant details. Sometimes she was simply wrong. But she was unerringly on target where it counted most.

That information became the foundation of a *Chicago Tribune* series (Sept. 10-12) which detailed how overwhelmed and under-trained nurses kill and injure thousands of patients every year as hospitals sacrifice safety for an improved bottom line.

MAUDE is available as your source, too.

Officially, MAUDE is the Manufacturer and User Facility Device Experience database offered freely by the U.S. Food and Drug Administration (*www.fda.gov/cdrh/databases.html*). It is also available

Continued on page ten

Inside Uplink

This month's *Uplink* details some of the recent stories done on racial profiling. Dan Browning of the *Minneapolis Star-Tribune* discusses his work on linking booking, jail and sentencing data in one Minnesota county (see this page). Jim Adams of *The Courier-Journal* in Louisville used police "check sheets" to see if police were unfairly stopping minority drivers (see page two). And Amber Arellano and Victoria Turk of the *Detroit Free Press* examined Detroit's suburban police forces to see if black drivers were more likely to be stopped in white suburbs (see page four).

Education

Football is king in Georgia high schools, but are girls sports suffering? Mike Fish of *The Atlanta Journal-Constitution* embarked on extensive records collection to look into Title IX requirements in the state.

SEE PAGE SIX

Tech Tip

A TV newsroom's Intranet set up is detailed by Alan Cox of WCCO-TV in Minneapolis.

SEE PAGE FOURTEEN

Uplink

December 2000

Volume 12, Number 10
A newsletter of the National Institute for Computer-Assisted Reporting

EDITOR
Brant Houston

MANAGING EDITOR
Mary Jo Sylwester

ASSOCIATE EDITOR
Paul Monies

ART DIRECTOR
Kerrie Kirtland

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.

DIRECTOR OF PUBLICATIONS
Len Bruzzese

SUBSCRIPTION ADMINISTRATOR
John Green

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 e-mail to jgreen@nicar.org

RACIAL PROFILING

Check sheets

By Jim Adams
The Courier-Journal

When Louisville's mayor suggested last spring that some city police officers were stopping African-American drivers unfairly, he was giving voice to a suspicion — but he had no hard facts to support his belief.

We promptly went looking to see if there were any. And at first, we came up as empty as the mayor's assertion.

We were told, for example, that the Louisville Police Department does not enter any data on traffic citations into its computer system, instead sending all citations directly to the state courts system. The courts system does enter the data. But that turned out to be useless to us because of the impossibly complex search required to separate traffic offenses from other types of citations.

On the verge of abandoning the idea, we turned one last time to some helpful veterans inside the police department for guidance. Could we get a detailed listing of all traffic stops? Yes, we were told, every reported stop is stored in the department's computer-assisted dispatching (CAD) system.

On the verge of abandoning the idea, we turned one last time to some helpful veterans inside the police department for guidance.

Is the driver identified in those records in any way?

No, we were told.

Is there any record anywhere identifying drivers stopped?

Well, actually, there is one place, one officer said: The "check sheets."

"Check sheets" are handwritten logs kept by dispatchers. When an officer stops a car in Louisville and then radios information from the driver's license, a dispatcher writes that information down, including race, before checking to see if the license is valid and if there are any arrest warrants outstanding for the driver. The department retains those sheets for a year.

Matching information

The Louisville police CAD reports and the "check sheets," we learned, held two pieces of common information that would enable us to match a traffic stop with a license-and-warrant check: police car "unit number" and time of day.

For example, if the CAD report showed that unit 315 had a car stopped between 10 a.m. and 10:20 a.m. on a given date, and the "check sheet" for that date showed that unit 315 requested a license-and-warrant check on a white female at 10:07 a.m., we would have a match.

But even if we could develop this information, would it be meaningful to a community wondering about the attitudes within its police department? Experts in racial-profiling research we consulted thought it probably would be — not necessarily as proof that racial profiling does or does not exist, but at least as a worthwhile indicator.

This approach also would yield information about innocent drivers who were stopped but not cited, they said, making it in one way broader than studies drawn solely from records of arrests or citations.

Using the state's open-records act, we requested the CAD reports of all traffic stops and all "check sheets" for 30 separate days. We selected the 30 days randomly from a 12-month period using Statistical Package for the Social Sciences (SPSS).

The police department eventually handed us a six-inch-high stack of "check sheets," and another six-inch-high stack of computer printouts of traffic stops covering all our requested days. The city also was able to give us the CAD information covering 13 of our 30 requested days on a floppy disk.

Unfortunately, since then the police department has no longer decided that "check sheets" are public records.

Doing the legwork

The job of cross-referencing these two information banks was fundamentally a

Continued on page three

Continued from page two:

Check sheets

manual task, as tedious as a Florida recount — although it was considerably easier for the 13 days the city gave us on the floppy disk. Mark Schaver, *The Courier-Journal's* CAR director, transferred that information to an Excel spreadsheet, which enabled me to search the CAD records of traffic stops for individual patrol car unit numbers. That was far more efficient than my eyeball scan of paper for the other 17 days.

The entire job took me, working alone, about three weeks.

City officers made 6,211 traffic stops on those 30 days, and viewing the information in the most conservative way, we believed we had reliably identified the race of 1,647 of those drivers, about one-fourth.

In the end, we found that 44 percent of those "race-identified" drivers stopped and checked for warrants were African American, in a city where an estimated 27 percent of the driving-age population is black. Experts say that justifies a conclusion that blacks in Louisville were stopped and checked for warrants at a rate twice that of whites.

We also found that in the city's police district with the highest percentage of whites, 23 percent of the drivers stopped were African American — four-and-a-half times the district's white population. Which raised another question: Regardless of the residential population, what if 23 percent of the district's actual drivers are African Americans?

Counting traffic

The experts we consulted thought it was imperative that we confront that question by actually counting traffic in that district. Which raised yet another difficult question: Where, exactly, do we go to do that?

MapInfo software showed us the answer at a glance. In our sample of 1,647 "race-identified" stops, 267 were within that district — and one-third of those were along one major surface street that bisects the district. Mirroring the district numbers overall, 22 percent of our "race-identified" stops along that street were of African Americans.

We dispatched four teams of reporters to count traffic along a four-mile stretch of that street over five days during the hours when most of the stops in our sample occurred. The result: Of more than 3,800 drivers we observed, 7.6 percent were black — one-third the percentage of blacks stopped and checked for warrants in the district.

This was important confirming information, the experts said. It was a real-world observation that tended to reinforce a major question raised by the overall numbers.

Given the nuances inherent in racial-profiling research, these numbers as a whole proved nothing, but they were persuasive enough to raise a reasonable question, our experts said.

When an editor and I arrived in the police chief's office to interview him, the chief had a guest waiting with him — an assistant professor in justice administration from the University of Louisville.

The Louisville Police Department, however, didn't have any interest in addressing that question. When an editor and I arrived in the police chief's office to interview him, the chief had a guest waiting with him — an assistant professor in justice administration from the University of Louisville.

She proceeded to explain why our research was invalid. Her primary criticism concerned our sample. Rather than drawing 30 days at random from a year, she said, we should have drawn randomly from all 75,000 traffic stops made during that year. Ideally, she argued, we would have drawn 10 percent of them — about 7,500.

Our experts disagreed. So did another expert we called at the suggestion of the Louisville professor herself.

Jim Adams can be reached by e-mail at jim-adams@courier-journal.com.

The stories for *The Courier-Journal's* series are at *www.courier-journal.com*. To find the archive, pull down the site map window at the bottom of the main page.

The story is also available in the IRE Resource Center, story #16812.

RACIAL PROFILING

Race and tickets

By Amber Arellano and Victoria Turk
Detroit Free Press

African Americans have complained about racial profiling in Detroit's suburbs for generations. They say suburban police unfairly target them for tickets when they venture outside of predominantly black Detroit and into the metro area's mostly white suburbs.

No one had ever been able to prove — at least statistically — that Detroit's suburban police racially profile, however. To do it, the *Detroit Free Press* had to show a pattern of disproportionate ticketing by race — and in the process, take on a recalcitrant judge and 3,000 traffic tickets, among other challenges.

We chose to look at Harper Woods, a suburb that borders Detroit's east side, because of a good tip we had received. Plus, unlike most Detroit suburbs, Harper Woods cops track race on their tickets, making it possible to analyze their racial ticketing pattern.

The first question was how to do it. Harper Woods court data isn't computerized, so we had to manually sort the tickets. With piles of complex information to analyze, we knew we needed to use a database.

Getting the data

With our computer-assisted reporting expert Heather Newman's help, we set up Microsoft Excel spreadsheets with the categories we wanted to track: the driver's name and address, age, race, gender, location of ticket, infraction, time and date, car type and color, fine paid, and the officer's name. We decided 3,000 tickets was a good amount — about six months worth, a respectable sampling.

Beyond race, we wanted to see what the trend was in terms of location. That is, whether drivers coming from Detroit, or driving closer to Detroit, were more likely to be ticketed.

Then there was the question of access. By law, journalists can look at courthouse files without filing a FOIA. But we encountered numerous roadblocks.

They wouldn't allow us to photocopy the tickets to speed up the process, then one day the courthouse administrator kicked us out for talking to a court officer who happened to come in the room. The judge eventually conceded and allowed us back in, but with restrictions that included allowing only one person to look at files at the same time and requiring his workers to review our work. Later, he allowed us to bring in a portable copier. By then we were three weeks behind schedule.

Setting up the database

Meanwhile, Database Specialist Victoria Turk had been setting up the database queries based on a list of initial questions: the racial and gender breakdowns of people ticketed, the most common tickets for African Americans vs. whites, which intersections were most heavily ticketed, etc.

When Turk finally received the raw data, she had about 80 percent of the database queries set up. The results of Turk's ongoing analysis helped dictate some of the Arellano's reporting.

We often had mix-ups in the data, such as accidental duplicates or the wrong spellings of police officers. We also had to be careful when it came to the racial category. Officers check a driver's race based on their presumption, not by asking the individual. So when an officer didn't know, or put "O" for "Other" (they told us that sometimes happens with drivers of Hispanic or Arabic names or appearance, for example) we made sure they were counted that way, so we could track people by race as unambiguously as possible.

Most of the data inconsistencies resulted from the manual input process.

Four different individuals keyed in the data, and although they were working from the same categories and code list for offenses, there were inconsistencies where abbreviations were used for city names, spellings of street names, and virtually every data input item entered from the ticket.

These weren't necessarily typos; the ticketing officers themselves misspelled information or abbreviated creatively, or simply had poor penmanship. In retrospect, some of the data inconsistencies could have been avoided if clear and explicit directions had been given to our data keypunchers at the start of the project.

We also had some problems reaching our

Continued on page five

The Detroit Free Press' story, "When race adds up in traffic," is available online at *www.freep.com/news/metro/race1_20000601.htm*, or in the IRE Resource Center, story #16685.

Continued from page four:

Race and tickets

total sample goal of 3,000 tickets.

Either through the keypunch process, or in the process of being provided the paper tickets from the court, we somehow managed to get some duplicates in the data that needed to be eliminated.

If a ticket was inadvertently entered twice, it would affect our analysis to double count it. Once we eliminated the duplicates, we ended up having to return to the court to pick up a few more to meet our sample quota. This could have been avoided by keying in a few extra tickets, in case some turned out to be unusable for one reason or another.

Later, Turk saw that she could perform some calculations more easily in Excel. The calculated fields, as well as the original input data, were then exported to an Access database.

No one had ever been able to prove — at least statistically — that Detroit's suburban police racially profile, however. To do it, the *Detroit Free Press* had to show a pattern of disproportionate ticketing by race.

The queries performed tried to explore the role that race played in ticketing, if any. We looked at race differences for nearest intersection, offense, average dollar amount and total tickets.

Eventually, we looked at race differences by the ticketing officer. We looked at ticket dollar amounts by car model year to see if older cars were ticketed more because of defective equipment. We also took a look at the most heavily ticketed intersections, time of day, and model years to give our readers a feel for the way the Harper Woods police patrolled.

Traffic study

We had one more hurdle: whether black drivers were being disproportionately ticketed. To find this, we had to figure out the number of black drivers on Harper Woods' roads and to compare it to the ticketing pattern.

To do that, we conducted a traffic study, similar to one done in New Jersey by investigators who looked into the state police's ticketing patterns. Arellano sat at Harper Woods' top ten most ticketed intersections and documented the race of almost 8,000 drivers. A profiling expert, a Temple University professor, advised us how to do it.

It took a few hours every day for about five days. If we couldn't tell what the driver's race was — because of a dirty windshield, or someone whose race wasn't obvious — we didn't count them.

Finally, after three months of off-and-on work, we published the package June 1. The lead 1A story showed that Harper Woods Police are overwhelmingly and disproportionately ticketing African Americans.

Suggestions

If you're thinking about a similar database investigation, here are our suggestions:

- Plan ahead with your editors and coworkers, especially when it comes to setting up the database, and give yourself extra time.
- If you run into access problems, be diplomatic but firm. In this case, only the threat of a good fight would budge the court judge. And of course, know the law.
- During a difficult project where you and others are encountering many hurdles, encourage and reward people throughout the process. Good guacamole, brownies —whatever you make well — and small gifts are appreciated and deserved.
- Fact check the database results as many times as necessary, and by different people.
- Avoid Harper Woods District Court.

Amber Arellano, staff writer, can be reached at arellano@freepress.com
Victoria Turk, database specialist, can be reached at turk@freepress.com

Tipsheets

Tipsheet #1318. The *Lexington Herald-Leader* looked at the racial disparity in traffic tickets issued by the Lexington police (2000, Lexington).

Tipsheet #1320. Offers tips on the differences between racial profiling and racial bias (2000, Lexington).

NATIONAL CAR CONFERENCE

Data and privacy

By Paul Monies
IRE and NICAR

While privacy on the Internet has garnered a lot of attention these days, the whole issue of privacy needs to be explored more by the media, especially in light of some conflicting views by the public.

Specifically, polls suggest that while nine out of 10 Americans want more privacy, many people willingly surrender personal information to large corporations in return for a discount on groceries or for a few extra air miles.

"We talk out of both sides of our mouths when it comes to privacy," said Charles Davis of the University of Missouri's Freedom of Information Center. Davis shared his views at a panel on "New Media Ethics and Privacy" at the 2000 National Computer-Assisted Reporting Conference in Lexington, Ky., held in September. He was joined by Rebecca Daugherty, director of the Freedom of Information Service Center at the Reporters Committee for Freedom of the Press, Robert O'Harrow of *The Washington Post* and Dan Gillmor of the *San Jose Mercury News*.

Consumers may gain extra goodies by sharing private information, but marketers gain much more comprehensive data on consumer behavior. At the same time, however, journalists have a tough time convincing the general public that personal information, laid out in government databases, should be open and accessible to all.

Davis said journalists should remind their readers of the difference between data collected by corporations and data collected by government agencies. "We have to tell the public the way we use data is different from how Procter & Gamble uses it," Davis said.

However, O'Harrow admits that journalists haven't even laid enough groundwork to educate readers on why public records are important, let alone private ones. Davis said that critics want to close off information being posted by government agencies on Web sites because of privacy concerns.

"Unfortunately, the government is using scare tactics to close off public records," Gillmor said.

In one more example, a member of the Lexington audience told of the unintended morphing of privacy legislation in Colorado. The 1994 Driver's Protection Act prohibits state DMVs from reselling certain driver's license data, but since local police departments in Colorado collect some of the same information, they conclude they are arms of the DMV.

"That's a great example of what's going wrong with privacy," Gillmor said.

Noted Davis: "Privacy is becoming an industry. It's not just a topic anymore. There's a whole cottage industry emerging to protect personal privacy."

O'Harrow, who covers privacy issues for *The Post*, said that the fledgling privacy beat is much like the early days of civil rights reporting and environmental reporting.

Some issues of concern on the privacy beat include when companies legitimately get public records and turn around to resell and append it with other information, and the lack of individual control over information once it is out there.

For example, drug companies can buy prescription information from pharmacists. O'Harrow said that while most people think medical records are confidential, in truth they are among the least protected records in America. Specifically, HMOs, banks and security firms all have access to insurance records, which ultimately lead to medical records.

"When bad medical things are happening, public records are important," Daugherty said. "The First Amendment should be a huge defense for that."

But Davis says that there are various uses for data, and it's not just journalists who can claim a public interest. Since researchers and historians all use public records and government data, Gillmor said he sees a problem with setting journalists apart from the general public. "What are the definitions of a journalist in the Internet era?" Gillmor asked. "Serious problems arise when government defines what a journalist is."

Noted O'Harrow: "Privacy is a great story. These stories have a very real impact on very real people. There's room for explaining the issue in real ways."

Paul Monies can be reached by e-mail at paulm@nicar.org

Web links

The Freedom of Information Center at the University of Missouri is a reference and research library that has more than one million articles on access to governmental information. Its Web site is *http://web.missouri.edu/~foiwww/*

The Reporters Committee for Freedom of the Press *www.rcfp.org*

Tipsheets

Tipsheet #1295. This tipsheet provides contact information for organizations that deal with privacy issues (2000, Lexington).

Tipsheet #1294. Provides information on what to do when public officials hide behind privacy protection (2000, Lexington).

Tipsheet #1292. This tipsheet discusses new Federal Trade Commission rules regarding financial information (2000, Lexington)

Continued from page six:

Equity

application forms to determine how the money was allocated by sport and gender.

The newspaper also obtained from the U.S. Department of Education an electronic copy of every civil rights case filed since 1985 involving an education institution. The information was split into two databases and coded. With some programming, the databases were merged, the codes were turned into English and the results were posted on *ajc.com* in a searchable form.

Separate mail surveys were sent to all (more than 3,500) varsity high school coaches in Georgia, to 323 principals, 180 superintendents and to the 39 members of the Georgia High School Association, asking about budgets, participation rates, their teaching and coaching backgrounds, athletic facilities and gender issues.

Ninety-three percent of superintendents, 92 percent of executive committee members, 67 percent of the principals and more than 26 percent of the coaches participated. For some of the surveying, the newspaper worked in cooperation with sports researchers at Georgia State University.

The newspaper also developed a survey form that was sent to each of the state high school and activities associations in the country, and it polled the 46 private schools in Georgia that are members of the Georgia High School Association.

Finally, the newspaper obtained access to the Georgia High School Association's financial records for the past five years, and it obtained in electronic form the GHSA's 1998-1999 athletic participation survey for every sport at every high school, broken down by gender. The association initially resisted providing the data, but complied after its attorney concluded the GHSA would lose an open-records case if it went to court. The data was converted to a searchable form and included among the offerings on *ajc.com.*

Mike Fish can be reached by e-mail at mfish@ajc.com.

Continued from page seven:

Social science

tallized after several mass shootings in 1999, including Columbine High School and the Atlanta day trader shootings.

"There was a notion in people's heads that there's a strong sense of concern about this 'new' violence," he said. "We set out to explain it and it started as a regular, narrative newspaper story but we decided to go deeper."

Fessenden said the biggest problem the team faced was defining the terms used for researching rampage killings. *The Times* used its news research department as well as the AP's state wires to gather information.

"News databases are nowhere near as complete as you think," he said. "There's lots of holes in coverage and big differences between Dow Jones and Lexis-Nexis."

To manage the huge amounts of information, the team turned to NewsEngin and Lotus databases for data entry and analysis.

Fessenden said the series faced hurdles because there was a lack of quality social science research on rampage killings to give context to their findings. The series used the FBI's Supplemental Homicide Reports for details on each killing.

To balance news databases such as Lexis-Nexis and Dow Jones, which have limited information prior to the mid-1980s, Meyer suggested getting back to people.

"Identify that 'old guy' in the newsroom and ask them what they remember," he said. "Ask the librarians and compare it with what you have for the 1980s to give you some way to estimate the error rate."

In response, Fessenden said he was limited by the data they had access to.

"If we had gone further, it would have been a geometric increase in work," he said. "But we didn't base any data over time on our database—that was from the Supplemental Homicide Report."

While both Reeves and Fessenden admitted the stories were starting points, Meyer congratulated them for the work and their ability to convince management that the projects were worth doing.

"This is social science because it tells you what they did and how they did it," Meyer said.

Paul Monies can be reached by e-mail at paulm@nicar.org

Tapes of panels from the National CAR Conference in Lexington, including those discussed in this issue, are available from Sound Images, (303) 649-1811.

More information, including tape numbers for each panel, is available at *www.ire.org/training/audio.html*

When journalism meets social research methods and a look at two case studies: Rampage killer and priests with AIDS: Tape #CAR00-45.

New Media and CAR: Privacy, ethics and online data: Tape #CAR00-48

Continued from page one:

Nursing

from NICAR. With more than one million records, the database tracks adverse events involving medical devices. With a twist in thinking, it can also be used to track cases involving human error.

Human errors

Amidst the dozens of fields are two 1,000-character text fields, which provide mini-narratives of each incident. In thousands of cases — plucked out through Access queries — nurses were believed responsible for fundamental errors.

How many patients have been accidentally overdosed through the use of infusion pumps? How many patients were fatally trapped in their bed rails? How many patients connected to life-saving machinery died unnoticed in their hospital beds?

MAUDE was just one of a dozen major sources for this project. Despite the narratives, MAUDE typically offers vague, inconclusive information. There are no patient or hospital identifiers. Fortunately, there are companion paper records that can yield more details — and I pulled many thousands of reports — but MAUDE is more of a pointer than a destination.

I exported Access query files to Excel and began building a custom database beginning with the sketchiest of clues gleaned from MAUDE. "Blanks" were filled through the use of a dozen state and federal public record sources, from federal healthcare investigative reports to court records.

Other data

My custom database became a giant jigsaw puzzle. MAUDE helped define a border, but other databases and paper records filled in the picture.

The lack of definition to MAUDE was symbolic of the project's genesis. I launched into this subject with little more than a realization that the nursing profession was a largely unexamined, essential component of the U.S. healthcare system.

There were no signature lawsuits or events. The *Tribune* encourages reporters to scout out general themes. I initially looked at workplace safety issues surrounding the way medications are labeled, but I quickly settled into the challenge of quantifying nursing-related errors.

Exploiting a back-door FOIA strategy, I obtained five years of nurse disciplinary records spanning every state and U.S. territory. The National Council of State Boards of Nursing *(www.ncsbn.org)* compiles this non-public database. However, paper copies of the database are regularly sent to state licensing agencies, which turns the privately compiled data into a public record. Illinois reluctantly turned over the massive files, which were scanned into Access.

Violations receive numerical tags as well as descriptions, such as drug offense, medical error or felony conviction. After identifying key cases, I filed additional FOIAs for full investigative files from each state. These records can provide a startling picture into how patients are killed and injured.

Records revealed that some states withhold information about death or injury from public files as part of a plea bargaining process with nurses who agree not to contest administrative sanctions.

For an airtight picture of Illinois, I computerized 10 years worth of disciplinary records after spending two weeks combing through musty file cabinets located in the bowels of a state licensing agency. Many states offer limited records online through boards of nursing agencies *(www.ncsbn.org/files/boards/boardswebsites.asp)*.

MAUDE coupled with disciplinary records began to provide a clearer, provable picture of medical errors linked to nursing. But a Lexis-Nexis search of lawsuits provided more cases and filled more blanks.

Hospital data

Another gold mine of solid information is found in federal investigative reports by the Health Care Financing Administration, part of the U.S. Department of Health and Human Services.

HCFA provided computer searches from their extensive database. For example, I wanted lists of every hospital that had received a routine annual survey or had been the focus of a complaint investigation.

In most states, HCFA investigations are conducted by state public health officials who

Continued on page eleven

Read more about Berens' project in the May/June 2001 issue of the *IRE Journal*.

The MAUDE database is available from the NICAR Database Library. More information, including sample slices and record layout, is available at *www.ire.org/datalibrary/databases/maude/*

It is also available from the U.S. Food and Drug Administration, *www.fda.gov/cdrh/databases.html*

The *Tribune*'s story can be viewed online at *http://chicagotribune.com/news/nationworld/article/0,2669,2-46844,FF.html*

A hard copy will soon be available in the IRE Resource Center.

Continued from page ten:

Nursing

contract with the federal government. Reports are kept by HCFA and the state. FOIA reports from the state; it's much quicker. Federal law designates these files as public records. Even if there is a state law banning disclosure of hospital records — as in Illinois — federal public record laws supercede state law.

A handful of states post information on the Web. The best is Colorado's Department of Public Health and Environment *(www.cdphe.state.co.us/hf/hfd.asp)*, which provides searchable access and mini-narratives involving hospital investigations. Read these fascinating files for insights on what can happen inside a hospital.

Keep in mind that most hospitals maintain private databases. Some facilities turned over their data for inclusion in my custom database. For example, one hospital acknowledged dozens of medical errors monthly.

With the data, the *Tribune* unflinchingly reported that a majority of hospitals nationally are quietly ... creating a harried work environment that often compromises patient safety.

The *Tribune* also created custom databases to track staffing levels and other criteria, such as how many nonprofit hospitals had been purchased by for-profit companies. The Volunteer Trustees *(www.volunteertrustees.org/index.html)* — a national organization of not-for-profit hospital and health system governing boards — maintains a database of conversions.

Likewise, dozens of healthcare organizations have accumulated databases of information. Many will share portions of their records, which can provide pointers to court cases or state and federal documents.

The Institute for Safe Medication Practices *(www.ismp.org)* offers monthly alert bulletins about the most serious trends in medical errors. These bulletins provide informational signposts to specific cases.

For example, Pennsylvania-based ECRI research laboratories *(www.ecri.org)* have tracked a silent killer called free flow, which is the uncontrolled, sudden gush of medication through an infusion pump.

ECRI has a massive database collection as well as an impressive medical device library. By using their data and research — along with MAUDE and court cases, etc. — I was able to establish that about 150,000 infusion pumps in use today are capable of free flow, resulting in dozens of preventable deaths.

Results

The power of the *Tribune's* database hit home when an FDA official declared that the *Tribune* had found previously unknown deaths linked to free flow. The FDA requested a copy of our data.

The *Tribune* denied the request since the data was drawn from so many places, often confidential medical and federal records.

The *Tribune's* three-day series highlights that sophisticated, influential CAR can be crafted with the basic tools — Access, Excel and most any desktop computer.

Even the custom database was simple in concept. Fields included the standard fare of dates and places, but also included drug names to track patterns of medication errors and a "source" field, which detailed the specific record used to track the death or injury.

With the data, the *Tribune* unflinchingly reported that a majority of hospitals nationally are quietly eliminating or supplanting the role of their best-trained, highest-paid nurses, creating a harried work environment that often compromises patient safety.

In early December, the nation's largest private healthcare regulator, the Joint Commission on Accreditation of Healthcare Organizations, launched a nationwide crackdown on the use of infusion pumps capable of free flow, as outlined in my series.

Mike Berens can be reached by e-mail at mberens@tribune.com

Web sites:

Boards of nursing agencies: *www.ncsbn.org/files/boards/boardswebsites.asp*

The National Council of State Boards of Nursing: *www.ncsbn.org*

The Volunteer Trustees *www.volunteertrustees.org/index.html*

The Institute for Safe Medication Practices *www.ismp.org*

HealthCare Financing Administration *www.hcfa.gov*

The Star-Tribune's story is available at *www.startribune.com/news/metro/profiling/.*

It's also available from the IRE Resource Center, #16752. Call (573) 882-3364 for more information.

Continued from page one:

Race factor

conviction or other types of less serious legal mechanisms used to resolve some cases without a conviction.

We analyzed only those charges that were identified by the county as the "main charge" when the suspects were booked; we ignored subordinate charges — including warrants — which often get dropped in plea bargains or relate to other minor charges, especially traffic and parking citations.

Although our analysis examined all races and Hispanics, we focused on blacks and whites because their numbers are larger, and because we believed the races were less likely to be misidentified.

Hennepin County is the one Minnesota jurisdiction using a computer system that allows easy tracking of an offender's case from arrest to adjudication.

Our analysis examined the bookings in two groups: those originating in Minneapolis and those originating in all other communities within the county combined. This helped minimize the effects of chance or the effects of a particularly aggressive police officer in some of the smaller communities in the county.

City and county findings

The results for Minneapolis and the balance of the county were similar. Whites were less likely than blacks to be sentenced for most felonies. But blacks were less likely to be sentenced for minor or petty charges than were whites. This pattern was true in nearly two-thirds of gross misdemeanors, and two-thirds to three-fourths of misdemeanors.

Those crimes that typically require more police discretion were less likely to result in a sentence for blacks. Indeed, some crimes, such as lurking with intent to commit prostitution or to sell drugs, rarely resulted in sentences for anyone, black or white. But such charges are used far more often for blacks than for whites. Such disparities were more pronounced in Minneapolis than in the balance of the county.

To calculate booking rates, we had to estimate the annual Minneapolis population by race and by age, since no official numbers exist. We used census data and annual population estimates for the county and the state. We derived the city's population figures by adjusting them to the proportions and trends in the census data and other official sources.

The population figures are a rough estimate. But after comparing them to the estimates from an economic forecasting group, we concluded our figures were reasonable and conservative. Besides, the gap between black and white booking rates was so wide that even if our population estimates were slightly off the mark, it would not have appreciably changed the findings.

After adjusting for population differences, blacks were arrested more often for most crime categories than were whites. And for crimes requiring police discretion, they were generally less likely than whites to be sentenced.

While these results do not demonstrate "racial profiling" by police, they do support claims that bias often results in unfair treatment for many blacks. Our analysis should not be misinterpreted as an assessment of the police motivation for the arrests, however. We could not say whether police stopped someone on suspicion of driving without a license because he was black, or arrested someone for lurking because he was belligerent or simply because he broke the law. But we could measure the outcome of the arrests.

If sentences were a measure of the quality of the arrests, then police were less effective when arresting blacks on minor crimes. We computed one cost of that inefficiency by multiplying the number of arrests per crime category that failed to result in convictions by the average booking costs. The human cost — measured in terms of resentment, public humiliation and private suffering — is incalculable.

Dan Browning can be reached by e-mail at dbrowning@startribune.com.

Tread carefully

A question posted to NICAR-L in March asking for warnings about interpreting race and traffic stops prompted a flood of discussion. Answers to this question varied, but the general consensus was that analyzing traffic stop data for any indication of race discrimination is a thorny proposition.

This is a condensed version of that discussion. To see the full thread, search the NICAR-L archives at *www.ire.org/datalibrary/nicarl.html*

Terminology

Distinguish between traffic stops and tickets. Robert Bennicasa, of Gannett News Service, wrote: "But I'd be wary of inferring that tickets were 'legitimate' or not, or constituted harassment, on the basis of conviction rates between two groups. The groups might tend to have different reasons for and rates of challenging tickets, with some factors being lack of time, lack of faith in the system, ability to pay the fine, ability to get time off work to go to court, etc. Also, one group might have a lower threshold for challenging tickets, pleading not guilty to 'legitimate' tickets more frequently, and the other group might only challenge those that are more defensible in court."

Demographics

Determining the demographics of highway users is necessary in order to put the data into context. If there is a higher proportion of tickets given to minorities, that would not be unusual if the particular geographical area in question is predominately minority.

Determining this, however, could be time consuming. The best suggestion offered was to collect the information yourself by sitting alongside a roadway for a few days.

Steve Doig, of Arizona State University and formerly of *The Miami Herald,* did just that for a traffic ticket study of a one-main-street small town in South Florida. Another reporter and Doig spent a day counting heads of drivers at the main intersection.

"The biggest problem with racial profiling stories is data tends to be incomplete and unless you're willing to do the field collection that Steve Doig suggests, it's hard to prove the police's excuses are wrong," wrote Drew Sullivan, formerly of *The Tennesseean.* Sullivan wrote a racial profiling story in February 2000 (*www.tennesean.com/sii/00/02/20/my profile 20.shtml*).

To figure out who is on the road, Sullivan said you can't look at Census numbers because blacks tend to be younger and the number of black drivers is less than their population as a whole. If you look at drivers in a city or county, you're not getting the commuters or people driving on the Interstate. Then there's the problem of whether or not to include drivers with suspended/revoked or cancelled licenses.

Minor offenses

Beware that socio-economics may play a role, i.e. tickets for minor offenses such as equipment problems, making it inaccurately appear race is a factor.

Frank Kummer, of the *Courier-Post* in Cherry Hill, N.J., wrote: "I would guess that if the city in question is mostly minority, the tickets would reflect that. So, a story saying more blacks are ticketed might be incorrect. Secondly, tickets for so-called small offenses can be misinterpreted as simple harassment. For example, you see a ticket issued to a black motorist for 'improper use of horn.' You might assume then that he was stopped simply for being black. You might be right. But, you also don't know if the cop actually 'downgraded' from a bigger offense to give a driver a break. This happens quite frequently.

"I think it's easy for us not to like cops and assume the worst. Indeed, that's easy to do so given our experiences. But I also believe you should approach it objectively."

A few last bits of advice: Study which officers gave the tickets, and study the types of offenses.

And one more from Drew Sullivan: "Be careful how you use the term racial profiling as well. It has a specific meaning – using race as part of a profile of a person likely to commit a certain type of crime. I'm guessing it came about when drug interdiction programs set up specific profiles of drug couriers. If race was part of that profile, it was racial profiling and the ACLU got mad. The point is the police are looking for a specific crime – usually drugs. It's not the same as a cop giving a ticket to a black motorist because the motorist is black. That's more akin to racism than racial profiling."

Stories

For more on racial profiling, take a look at these stories available from the IRE Resource Center, (573) 882-3364.

Story #16331. *The Seattle Times* takes a look at traffic citations and race in Seattle over a five year period.

Story #16502. The *Village Voice* goes beyond the typical "numbers" story and profiles black and white police officers as well as victims of profiling.

Story #16874. *The New York Times* reports that senior officers knew of racial profiling by police in New Jersey at least as far back as 1996.

Story #16661. This Newark *Star-Ledger* story was one of the first on the widely reported problem of racial profiling in New Jersey.

Story #15652. *Emerge* magazine looks at the "Driving while black" phenomenon.

TECH TIP

Intranet set up

By Alan Cox
WCCO-TV, Minneapolis

I knew our computer-assisted reporting was getting out of control on "The Day of Eleven Names."

We had what I suspect is the typical experience of newsrooms that stick with CAR for a few years. We collected a few gigabytes of data for investigative projects. The same data would be a blessing for daily news — if we could figure out how to access it within a day.

When CD recorders became affordable, I began pressing the data onto discs. We bought a second copy of the database program FoxPro and put it on a communal computer at the assignment desk. But almost everyone was put off by FoxPro's complexity (or maybe it was the length of the instructions I wrote: "Step 1: Use the mouse and click twice on the little picture of a fox..."). So lookup requests kept coming to me.

The worst day

I kept a count on what I remember as the worst day: in eleven separate requests, reporters brought me names to "check on your computer." I had to spend 20 minutes on each name, switching CDs, watching the query progress bar slowly creep along, printing out the results and delivering them to the reporters (which I came to think of as "putting the facts on a silver platter for the ingrates"). Soon, we had to back off our efforts to spread the use of CAR.

Today, technology has become cheap enough that the single-computer approach might work. Thirty-gigabyte hard drives are reasonably priced, eliminating the need to shuffle CDs. And a few months ago I saw a demonstration by the CAR specialist at one of our competitors. He created attractive search screens within FoxPro that are customized for major databases.

But I was a non-computer-assisted reporter long enough to know that you shouldn't expect reporters to abandon the telephone at their desks to look for something on a shared computer. And assignment people have enough to worry about without having to serve as librarians for the staff.

Eventually we experienced a wonderful coincidence: the station's budget for buying computers had a small surplus, and NICAR announced a seminar for creating newsroom Intranets. My news director gave me both.

There were two challenges. One was getting the new server to operate on the station's computer network. For this I had generous help from our information technology manager.

In fact, I'm not sure how generous it was. He used to be the engineer who ranour satellite truck, and after years of watching me do live shots, I don't think he was going to let me come near plugging a computer into our network.

The other challenge was the industrial-strength database program, Microsoft SQL Server. It works like FoxPro or Access, but has a number of devilish differences in the details. And, at least in the version we bought three years ago, it generates the most indecipherable error messages I've ever seen.

I was a non-computer-assisted reporter long enough to know that you shouldn't expect reporters to abandon the telephone at their desks to look for something on a shared computer.

What made it work for me was the program that serves as an intermediary, NewsEngin's UniQuery. Often praised in these pages, UniQuery creates search forms for the databases loaded on the server. It installs easily. It creates forms on Web pages, and by now almost every reporter has experience with that interface.

Using UniQuery

Best of all, UniQuery creates a unified

Continued on page fifteen

Tipsheets

For more on building and maintaining newsroom Intranets, check out these tipsheets, available from the IRE Resource Center, (573) 882-3364.

Tipsheet #895. NewsEngin's George Landau offers tips on getting started with an Intranet (1999, Boston).

Tipsheet #894. Offers tips on running an Intranet (1999, Boston).

Tipsheet #817. This includes useful information to include in a newsroom Intranet (1998, New Orleans).

Tipsheet #1328. Results of on an Internet survey on NICAR-L concerning newsroom Intranets (2000, Lexington).

Tipsheets are also available on IRE's Web site, *www.ire.org/resourcecenter/initial-search-tipsheets.html*

Continued form page fourteen: **Intranet**

search page so reporters can hit every database at once. We have 70 queries loaded on our Intranet now, and SQL Server is powerful enough to complete them in a half minute. There's a wonderful serendipity in finding extra details about a person in databases you'd never have time to check otherwise — a phone number on a hunting license or an employer in campaign contribution lists.

Once UniQuery was up and running, I ventured into creating my own Web-based search pages for other data, using training aids created by Tom Torok, now of *The New York Times*. I created profiles of public schools drawing data from three different sources. I listed city data from FBI crime reports.

Even simple Intranet features have their benefits. I got permission from the state to post the Minnesota criminal code (not that I fully accept why I need permission for that). It's copied straight from the state legislature's Web site. But the indexing built into the server's operating system is simpler to use and is up and running more often than the legislature's search engine.

We've even branched out to other features used by the entire station, like a Web form to send messages to our alphanumeric pagers. The most widely used features within the station are a form to fill out expense reports on a spreadsheet and our overnight TV ratings.

Lessons from the Webmaster

All of which means I've turned into a Webmaster. That's taught me lessons that may apply as we all try to create versions of our work for the Web as well as for broadcast.

First, not that a TV reporter should need to be reminded to be concise, but I thought the Intranet was supposed to be different — the more details crammed in, the better. Instead, I've been surprised to see how slow users are to "drill down," to click repeated links to get "deeper" into the data for more details. I've even watched people who seem reluctant to scroll down a single page to look for data. We're about to receive a new version of UniQuery that should simplify the look of the results.

I've learned not to be cavalier about letting people figure things out for themselves. One day I watched an assignment person struggle on the Intranet to calculate how old a person was from the date of birth. I smugly shouted out an answer in an arrogant "I got straight-As in arithmetic" tone, and then had to go back in a minute when I realized I'd figured it wrong, too. I've now added an "AgeToday" column to the query results for databases that include birth dates. Let SQL Server worry about the math.

And although I no longer have the need to tell people to click twice on the little fox, I have tried to make instructions clearer based on user feedback (often shouted at me across the newsroom when I've got my own deadline to worry about). I assembled a PowerPoint tutorial on the use of our Intranet that I've updated several times to reflect complaints and cautions. One is, "These databases were collected from government agencies, and governments can make mistakes." That may seem too rudimentary even for Journalism 101, but I think it's a helpful reminder for people impressed by the presumed power of computerized data.

I assembled a PowerPoint tutorial on the use of our Intranet that I've updated several times to reflect complaints and cautions.

But maybe I contribute to that image of power. I keep a running tally of the size of the Intranet on its homepage, something like the McDonald's hamburger signs. Right now it reads "Now featuring 41.9 million records."

Eleven names? Bring 'em on.

Alan Cox can be reached by e-mail at wccotv@mr.net.

Web links

For more on Intranet software, check out the following Web sites:

NewsEngin
www.newsengin.com.
Under "Free Tools," you'll find a link to Tom Torok's tools and handouts from training seminars.

FileMaker Pro
www.filemaker.com

SQL Server
www.microsoft.com/sql

ColdFusion
www.coldfusion.com/products/coldfusion

Bits, Bytes and Barks

Census Data

The big question these days is, "How will I get the Census 2000 immediately after it's released?" The first data, state-by-state redistricting data, will come out sometime in March. The Census Bureau plans to make it available in several ways, including on the Web, via FTP and CD-ROM.

IRE and NICAR and the Associated Press are both studying ways to get the information to journalists the same day it is released. Details have not been worked out yet. However, NICAR's plan will include producing a standard set of extracted data, offering the most immediately useful summary levels (like state, county, place and tract). These would be extracted into Excel files for each state and put on an FTP server accessible to subscribers.

Stay tuned to *Uplink* and the NICAR-L and IRE-L listservs for more details about the NICAR Census plan.

More information about Census 2000 for journalists is available at Steve Doig's Web site, *http://cronkite.pp.asu.edu/census.*

Access Tech Tip

Here's a question posted to the NICAR-L list recently with a common, but annoying, importing problem: "I'm trying to import fixed-width text into Access, using the import wizard, but before the process begins, an error message appears telling me that the first row on my records is too large. The record length is about 320 characters. The text I'm bringing in includes about 300,000 records. Where have I missed the boat?"

The problem was that the file either didn't have end-of-record characters, or had ones that Access doesn't recognize. As a result, Access was thinking the entire 300,000-plus-record file was one big record and it choked.

There are several ways to fix this problem:

1) If the file isn't too large, you can open it in Word or Wordpad and save it as MS-DOS text.

2) If you have FoxPro in your office, another option is to open it in FoxPro's text editor and saving it as a new file. FoxPro will automatically add the end-of-record characters.

3) There are many text editing programs available, including demo and shareware versions. These will allow you to add in the necessary characters. Two possibilities are EditPad, *www.jgsoft.com*, or UltraEdit32, *www.ultraedit.com.*

Campaign Finance Data

The NICAR Database Library offers a monthly subscription to federal campaign finance data from the Federal Elections Commission. Each month, you would receive four tables of data, including individual contributions, committee contributions and other transactions, a listing of candidates and a listing of committees.

The subscription costs $100 for news organizations with circulation under 50,000 or 50-200 market; $200 for circulation 50,000 to 100,000 or 25-50 market; and $225 for circulation over 100,000 or top 25 market.

More information about this data is available at *www.ire.org/datalibrary/databases* or to place an order call NICAR at (573) 884-7711.

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