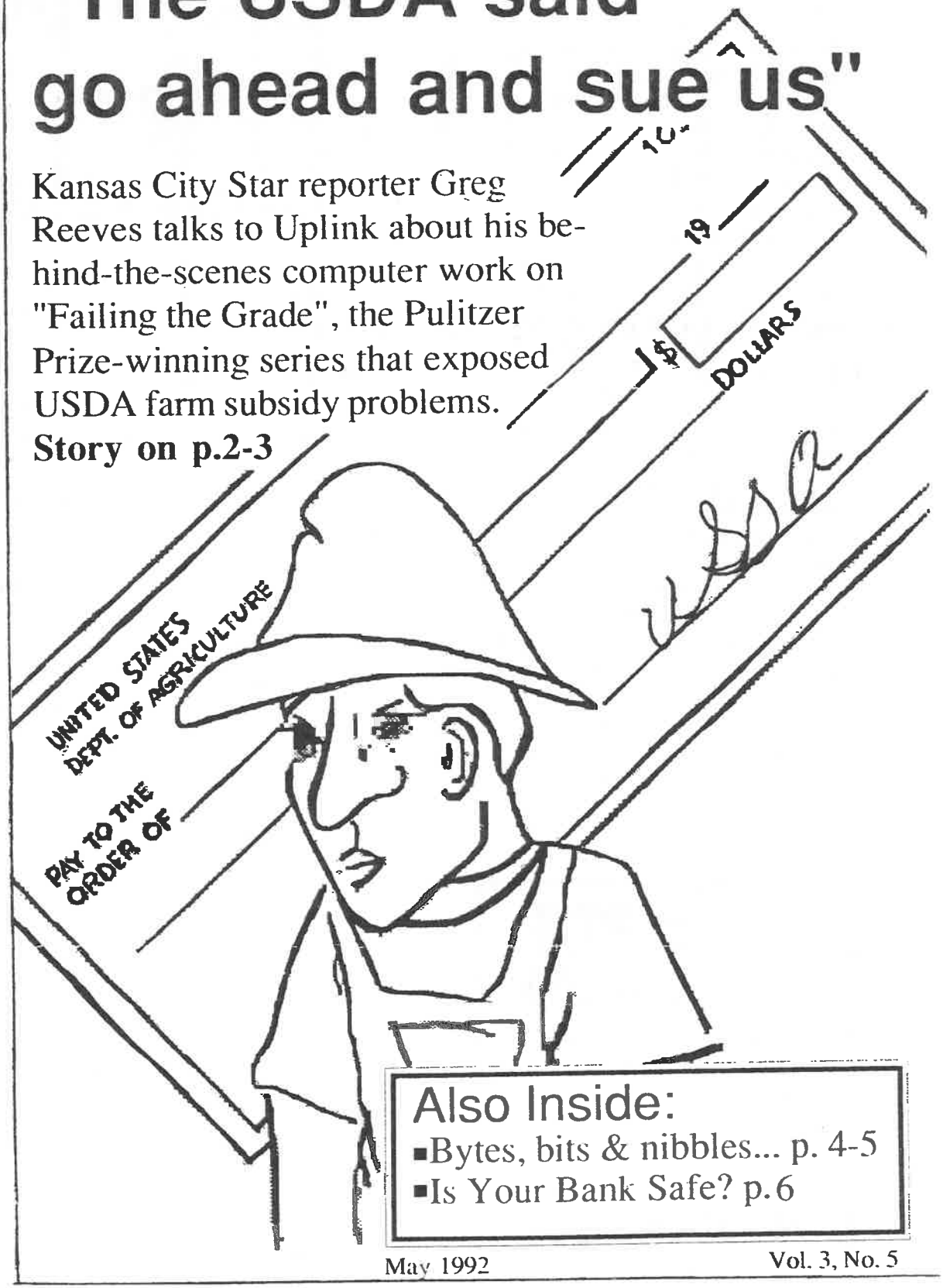
"The USDA said go ahead and sue us"

Kansas City Star reporter Greg Reeves talks to Uplink about his behind-the-scenes computer work on "Failing the Grade", the Pulitzer Prize-winning series that exposed USDA farm subsidy problems.

Story on p.2-3



Also Inside:

May 1992 Vol. 3, No. 5

The forum for computer-assisted reporting.

Dear Readers:

I would like to take this opportunity to introduce myself as the new editor of Uplink. I am taking over for Bob Jackson, who is still with MICAR, and can be reached at his old number. However, beginning with this issue, I will field any questions or comments concerning Uplink.

We also have added a new designer to our staff. Joyesha Bhattacharya created Uplink's new cover design, and in the coming months will be working on a complete makeover of the newsletter.

The purpose of the redesign is to build on Bob's foundation and make Uplink a more "user friendly" newsletter. Upcoming issues will feature more graphics and photos, plus greater opportunities for reader input. We want Uplink to be as useful a tool to you as your computer.

We appreciate continued correspondence from our readers: not only your successes, but your problems and pitfalls as you boldly go where few reporters have gone before.

-Please call us at (314) 882-0684 with any questions or comments.

— David Washburn
Editor

MICAR is interested in any information, ideas or stories related to computer-assisted reporting for future issues of UPLINK.
If you wish to contribute, please mail stories or ideas to:
120 Neff Hall
University of Missouri
Columbia, MO 65211
(314) 882-0684

Star reporter dodges USDA obstacles

By David Raziq
MICAR

"From the standpoint of computer-assisted reporting, this was a techno-nightmare," said Kansas City Star reporter Greg Reeves."It took three gigabites of space, the tapes blew up my Qualstar 1260 taped-rive and the number of records overwhelmed XDB." But Reeves isn't complaining, because the 'nightmare' he's describing is "Failing the Grade", a comprehesive and critical look at the United States Department of Agriculture that this April won his paper a Pulitzer Prize for national reporting. The effort of reporters Mike McGraw, Jeff Taylor, Reeves and editor Mark Ziemann, the seven-part series also won an Investigative Reporters and Editors Award earlier this year.

The Star's investigation began nearly two years ago as the result of previous stories by reporters McGraw and Taylor. One such story included the problem of "dirty grain": farmers and grain producers adding dirt to grain export to increase the weight and therefore the price of shipments. That and other chronic problems, says Reeves, convinced the Star that the USDA's troubles "went much deeper" than anyone realized and so the project began. Eleven months later, Reeves joined the investigation.

"For my part of it," he says, "I asked the USDA for the last five years in farm subsidy payments." The USDA pays out over ten billion dollars in subsidies and Reeves suspected he might find some irregularities to those payments. But Reeves' effort soon hit a snag when the government denied his request.

"We started throwing lawyers at each other," Reeves said, but "unlike other agencies that ultimately agree to give you the records, the USDA said go ahead and sue us. I think they denied us access simply because they had never given them to anybody before and because bureaucrats tend to believe its their infromation and not the public's."

Greg Reeves spent months trying to access USDA farm subsidy tapes.

But getting the tapes was the easy part.

Three months later a compromise was finally reached: The USDA would provide 8.2 million computer records of only 1990's payments. Those tapes included the check's serial number.the date, county, and state the check was issued for, and even an identifying number for the intended farmer but excluded the receiver's name and address.

However that was just the beginning of Reeve's problems, because with the tapes in hand, he now experienced difficulties in down-loading the information into his computer. "Our tapedrive could not read these tapes. The tape quality was marginal and the Qualstar 1260 is a piece of trash when it comes to reading marginal tapes."

Reeves says that although a recent upgrade has solved most of the Qualstar's reading problems, he felt discouraged at the time. Starting in July 1991, Reeves says he "worked nearly around the clock seven days a week to load these tapes to my PC... We failed."

So Reeves decided to install the nine tapes on the Star's Vax Mainframe processer using NineTrack Express for the download and then Datatrieve, the mainframe's database software, for the analysis.

"I wrote hundreds of programs to extract this data," says Reeves,"and in some cases ran each of them dozens of times" for a two-month period. Finally he created an analysis based on those farmers who got the most payments along with their county and state. But the investigation was still missing one thing: the names of the farmers or producers receiving the money.

So using the check's serial number, the date of issue, and purpose, Reeves began calling various county USDA offices to match the checks with their receivers.

As a result, one of the investigation's findings included the unchecked growth of what Reeves calls "Mississippi Christmas Trees."

"We've found that in ten states," he says, "there are more farmers getting farm subsidy payments than there are farms. What's happened is the farmer's have split up their farms under so many different ownerships in order to skirt the USDA's fifty-thousand dollar per farmer limitation." As a result, the Star's team found wives, children, cousins, uncles, and even family doctors and lawyers all listed as producers with the USDA and thus receiving subsidy checks.

Although Reeves depended on a mainframe for "Failing the Grade," he says he still primarily believes in PC-oriented research methods and has some recommendations for those organizations just starting their own computer-assisted programs.

"Go for a PC-based set up that takes care of your needs at least two years out: namely a 486 33 megahertz with a minimum two giga-byte harddrive," he says. In addition, Reeves recommends Foxpro, Paradox, or SAS database software for the analysis, NineTrack Express for downloading tapes, and a Pertech, Cipher, or Storagtech tape drive. "The whole set-up" says Reeves,"would be around twenty thousand dollars."■

Letters

Dear Uplink:

I got my April 1992 Uplink today and was intrigued by the reports of problems with XDB software. I have had some of the same problems, and I'll be waiting to hear what solutions arise.

I got XDB 2.30 in May 1991. We have not yet attempted any projects as complicated as those described by Robert Paynter, so I hadn't run into the problems with indexing, but I had noticed two other problems mentioned in the article: the software would give some sorts out of order (the decimal point example cited by Ed Foldessy of the Wall Street Journal is a perfect parallel), and its computations were not always trustworthy.

As an example of the former, I've run a fair number of queries on census data, and I frequently found that when I calculated percentages—say, ranking census tracts by percentages of minority residents—two percentages that were a point or two apart would occasionally appear out of order. I eliminated most of those problems by simply adding another float decimal place (F10 utility menu, then F5 user profile, then F6 format options, then F8 float decimal places), but that didn't get rid of them all, and even with version 2.41 I've had to keep a close on any calculations and oderings in which decimal places were significant. Perhaps I should've been less tolerant.

I ran into the other computational problems in working with a cut-down version of the 1989 Toxic Release Inventory data for North Carolina, I found that XDB 2.30 could rank them accurately statewide—it seemed to ignore any record with a toxics total exceeding seven digits, of which we had at least one. (This wasn't apparent from the data itself; fortunately, we had other TRI reports with which to compare some of our data and could tell right off that there was a problem.) However, version 2.41 totals the larger numbers accurately.

One other problem that 2.41 solved for us was using the PC's random access memory. We bought four megs of RAM with the machine, and in my relative ignorance of computers I'd assumed that if it was in the machine, it was being used. Meanwhile, when I tried to import census data into 400-field tables and kept getting the X10 error message—insufficient RAM—I just figured that four megs wasn't enough and built smaller tables. I didn't learn about basic versus extended memory until stumbling over Jack Warner's computer column from Cox News Service on the on the wire a couple of months ago—it explained what the problem was, although it didn't suggest anything to solve it that worked for us. Dave Stroble, our audiotext product manager and a PC fiend, futzed around with my AUTOEXEC.BAT and CONFIG.SYS files but couldn't solve the problem. Then we got XDB 2.41 and its extended memory manager. I don't know exactly how it works, but since I installed 2.41, I've had no more problems with totaling numbers or importing files with many fields. (I haven't had occasion to try 400, but I have gone as high as 288 with no problem. Not only that, 2.41 appears to import about 20 percent faster that 2.30.)

If XDB has figured out how to fix the ordering bug, I'd like to hear about it. Meanwhile, please keep us posted on other problems, even if it doesn't make for as compelling reading as does the Hartford Courants's serial killer!

Lex Alexander
Investigative/Database Reporter, Hartford Courant

Bytes, bits and nibbles..

Reporters and editors around the United States are warming up to the idea of integrating computers into journalism. And if the packed seminar on computer-assisted reporting at the **Investigative Reporters and Editors** conference in Atlanta is any indication, the coming years will see the concept catch fire in our nations newsrooms.

"This is not a luxury. This is not the future. This is a way to get information now," said **Bill Dedman,** I.R.E. Board Member and freelancer. Dedman, who is currently writing a book on computer-assisted reporting along with **MICAR** founder **Elliot Jaspin**, said the use of computers in investigative journalism is changing the way government and the media do business together.

The Freedom of Information Act and access laws are slowly beginning to facilitate the process for those who use this new technology.

"The general premise is that if you can get it on paper, you can get it on computer." But, Dedman said state and local governments are routinely putting roadblocks in the way of reporters who attempt to gain access to computer records.

For example:

The state of North Carolina's attempt to charge the **Raleigh News and Observer** more than $12 million for the state's driver license records, which are on nine-track magnetic tape. The state and the newspaper are currently in litigation to decide on a fair price for the records. "The federal government has spent billions of dollars inputting information on more than 32,000 mainframe computers," Dedman said. ""And it spends no money telling you which databases it has."

Despite the government's obstacles, reporters are turning out award-winning stories that would have been impossible to complete without computers.

The News & Observer has arguably one of the best computer-assisted programs in the country. With about half of their newsroom trained in analyzing computer databases, the newspaper churns out

News

computer-based stories on a daily basis.

Here's a taste of what the News & Observer has been working on the past few months:

*Computer analysis of North Carolina Marriage records for the past 20 years yielded an interesting story for reporter **Mary Miller.** Her analysis found that although Friday is the most popular day to get married at the Wake County Courthouse, the 13th is always a slow day. It seems that people are superstitious about the 13th, no matter what day of the week it falls on.

*News & Observer reporter **Carol Blaney** performed a computer analysis of the state's escheat and unclaimed property program and was able to detail how uncashed checks, stocks, forgotten accounts, jewelry and cash left at hospitals, banks and other businesses can be found. Her story, "Wake residents owed $305,579: Escheats office holding millions in unclaimed funds," explains how the state runs the gram and how the interest is used to fund the program and provide college scholarships.

*In February, reporters **Steve Riley** and **C.E. Yandle** performed a computer analysis of records from the labor department and the state office of the chief medical examiner. They found that because of Labor Department oversights, many on-the-job deaths remain uninvestigated. Their story, which ran last month, detailed seven cases of workers dying on the job. The deaths were never reported to all the correct authorities and never fully investigated. The News & Observer story stated that the Labor Department misses one worker death in four and the department relies on employers and the news media to notify them of worker deaths instead of checking in with medical examiners and taking a more proactive stance.

Also busy turning out computer-assisted stories are **Doug McInnis** and **Steve Stephens** of **The Columbus Dispatch.** The reporters recently completed a computer analysis of drug- crime records and conducted scores of interviews. The end result was a series entitled "Drug killings boom in just 16 months."

The series depicts the drive-by killings, crack killings and other drug killings that have shattered Columbus' image as a big city without big-city crime. The series includes graphics of the rising pace of drug-related killings and statistics revealing that they have become the leading type of homicide in Columbus.

Computers are being used more and more on breaking stories. For example, when a plane crashed in Dayton, a reporter using a computer discovered that the plane had previously had 32 serious maintenance problems.

Hal Straus from the **Atlanta-Journal Constitution** pointed out the trend toward using the computer to enhance solid beat reporting. Journal-Constitution reporters used a database to trace campaign contributions to Atlanta's governor and lt. governor. They found that the sponsor of a bill, allowing beverage companies to label just about anything short of sewage as spring water, had contributed heavily to the two campaigns.

"It took five minutes to search the database and get a really good story," Straus said.

IRE seminar

The widespread interest in computer assisted reporting has led IRE to establish its first all-day seminar on the subject at its national conference in Portland, Oregon. The conference will take place from June 11 to June 14, 1992. For more information telephone IRE at (314) 882-2042.

from around the nation

Easing the fear of a financial collapse

by Trebor Banstetter
MICAR

The bank of New England failed in early January of 1991. It was the biggest bank failure in the history of the United States, and the financial catastrophe caused great concern among New Jersey residents who kept their money in local banks and thrifts.

"The people of New Jersey were very nervous about the safety of their banks," said Dan Woods, business writer for the Record of Hackensack, N.J. The fear was so pervasive, Woods said, that a thrift in Patterson, N.J., had a "run" _ depositors withdrawing all of their savings _ because of a takeover by United Savings Bank. "There was poor communication, and people thought their money was going to disappear," he said.

"We wanted to give them some information on how safe their bank was."

Woods wanted to create a ranking of banks and thrifts that would assist consumers in choosing what financial institution was safest. He decided to use a publicly available database provided by Ferguson and Co.

The people of New Jersey were very nervous about the safety of their banks. We wanted to give the information on how safe their bank was.

-Dan Woods

The data was based on Federal Deposit Insurance Corp. call reports, which are detailed financial reports banks have to file each quarter, and are available on a CD-ROM disk. Obtaining the information from the disk and downloading it into a database program is very similar to using nine-track magnetic tapes.

The best way to determine a rating for a bank or thrift, Woods decided, was to use a formula comparing an institution's capital and its problem loans. "In a failed bank, the losses from the problem loans keep piling up and eating away the capital," he said.

Each bank's capital and problem loans was divided by the bank's assets to produce a percentage. The capital percentage gave an indicator of how much the bank had saved over the years to act as "a cushion of last resort." The problem loans indicated how good the bank was at making loans.

The capital percentage and the problem loan percentage were given equal weight in the ranking, and were converted into the standard deviation to equalize the numbers. Then, each bank had a ranking between one and five.

"After the story ran, consumers were wildly enthusiastic," Woods said. "This gave them a way to judge the basic safety of their bank."

One of the biggest concerns Woods had was making sure the story presented the data accurately and fairly. "We were very conscious of the fact that people would use these ratings to move their money, so we checked the data and methodology to the extreme," he said. He also pointed out banks and thrifts that had low rankings but good recovery programs.

Woods made sure to include information on deposit insurance, so readers understood their savings were insured by the federal government if under $100,000. "We didn't want to cause hysteria," he said.

In 1992, Woods did a second ranking, this time including several significant out-of-state banks. "Again, the response was overwhelming."

The stories have won several awards, including a third prize in the enterprise category from the New Jersey Press Association. They also won a C.T. Rutgers award for business reporting.

Woods encourages others to try similar rankings of financial institutions in their areas, but warns that extreme caution and care must be used. "If you print the wrong data, it can be devastating," he said.

But, he added, "It's definitely worth the effort. I have received calls from senior citizens who told me they would sleep a little easier."■