

At Your Readers' Fingertips

Searchable databases
of information relevant
to users' lives
increase Web traffic

by Heidi Ernst
Illustration by James Turner

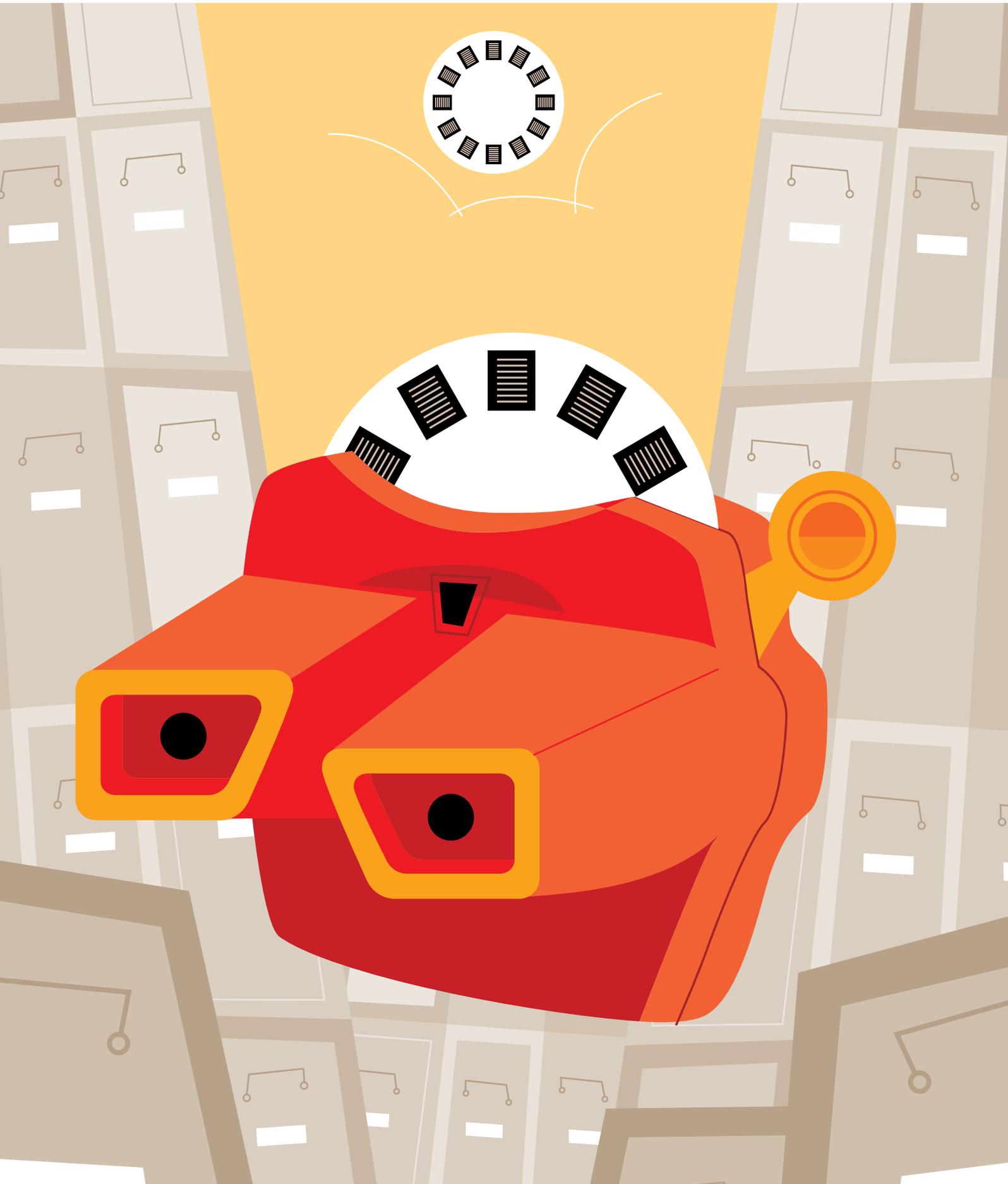
On May 1, The Gazette in Cedar Rapids, Iowa, ran a front-page story about the first-time release of a list of arrest warrants the sheriff's office hadn't been able to serve, dating back to 1965.

On that same day, the paper introduced its readers to Gazette Data Central (<http://gazetteonline.com/section/datacentral>), a clearinghouse of searchable databases, including a listing of the Linn County outstanding arrest warrants. Usage numbers were themselves arresting. On the first day, the warrants database had 117,000 page views; on the second, 88,000. That was impressive, considering the population of the area covered by the list is only 202,000.

More often, newspapers are increasing traffic to their Web sites and, in some cases, are selling more single copies of their

Nutshell

- ▶ Searchable databases are allowing newspapers to interact more with readers and create a new synergy between Web and print.
- ▶ Local investigative reporting is getting a boost.
- ▶ Fear of technology isn't a reason to sit on the sidelines; hire someone who is trained to create multimedia projects and databases.



Tips for Creating a Database

WHETHER YOUR newspaper is already posting online interactive content or just gearing up to, consider these ideas from representatives of newspapers with a portfolio of databases up and running.

▶ “Prioritize based on the needs of readers. Stop and think in a really practical way: How will readers want to use this? And create templates to apply to other databases so you’re not recreating the wheel.”

- Carolyn Washburn, editor and vice president of news, The Des Moines Register

▶ “Reaching out in partnership to the rest of the building is important. The marketing and circulation departments have a lot of information.”

- Bob Stover, managing editor, Florida Today in Melbourne

▶ “Start off with something small... you can launch with a few functionalities and leave more difficult functions for later. Then watch what readers are searching for, and let those answers direct your next phase in development.”

- Jennifer Ward, interactive media editor, FresnoBee.com

▶ “Get people in the newsroom to understand how important the integrity of the data is, even if you have to delay a project or put more people on it. Once you lose credibility, you lose your audience.”

- Traci Bauer, managing editor for multimedia and innovation, Democrat and Chronicle in Rochester, N.Y.

print editions by creating or tailoring information and posting it online as searchable databases. Like the Gazette, many are taking public records—such as property taxes and death notices—or collecting their own watchdog or lifestyle information, and generating online catalogs of searchable data.

How newspapers get ideas, transfer documents into searchable records and hire technical staff varies. The commonalities, however, include starting with data relevant to readers’ lives. Chuck Peters, president and chief executive officer of the Gazette Co. in Cedar Rapids, says newspapers are just beginning to grasp how to do that.

“More than specializing in the Web, it’s thinking about how to communicate knowledge rather than just stories,” he says. “It starts with people thinking differently about ways to interact with the community. We’re just starting to have the conversation.”

Catching the Local News

A newspaper’s top new-media strategies need to be “owning local news, generating hyperlocal content, and using database-driven coverage,” said Rob Curley, vice president of product development for Washingtonpost. Newsweek Interactive in Arlington, Va., at a content-management session at NEXPO® 2007 in Orlando.

Those concepts powered the development of popular multimedia projects he has overseen at various newspapers. They also apply to online databases created by other newspapers, including a searchable database about speeding, created by The News & Observer in Raleigh, N.C. ([\[newsobserver.com/news/growth/traffic/speed\]\(http://newsobserver.com/news/growth/traffic/speed\)\), and an interactive map and database about biofuels in Iowa, developed by The Des Moines Register \(\[www.desmoinesregister.com/biofuels\]\(http://www.desmoinesregister.com/biofuels\)\).](http://www.</p>
</div>
<div data-bbox=)

The same concepts propelled Gannett Co. in McLean, Va., to create Information Centers at each of its 84 local U.S. newspapers last year. The centers are organized around seven information-gathering areas, or desks, including the Data Desk.

“We believe databases have expansive potential for elevated public-service coverage,” says Jennifer Carroll, vice president of new media content for Gannett. “We have the tools not only to break stories based on data analysis but to give readers the ability to search and analyze in powerful ways.”

The company’s methodical strategy has led to powerful milestones at Gannett’s newspapers. The Democrat and Chronicle in Rochester, N.Y., for example, set a single-copy sales record with a “multimedia-first” police-salary database and had similar successes with property-tax and groundwater-pollution databases—all launched online on a Thursday (www.democratandchronicle.com/rocdocs) with an expanded print story the following Sunday. The paper also is doing more investigative work than ever, says Karen Magnuson, editor and vice president of news.

Other Gannett newspapers are finding online databases to be a powerful tool for reaching readers. Terry Eberle, executive editor of Florida Today in Melbourne, says his paper is “sharing more First Amendment material with readers on more topics than ever before.” And the Asbury Park Press in



“We believe databases have expansive potential for elevated public-service coverage.”

— JENNIFER CARROLL, GANNETT CO.

Neptune, N.J., has averaged 5 million to 9 million page views a month since last December, when it launched Data Universe (www.app.com/apps/pbcs.dll/section?Category=DATA), searchable databases of public records such as criminal convictions, crime reports, property assessments and public employee salaries.

Staffing Up

“This isn’t about picking just any information we can and posting it,” cautions Mark Bowden, editor of the Gazette, whose

Gazette Data Central now posts almost a dozen searchable databases, from a local dining guide and a summer events calendar to a crime map (see sidebar below). “It’s about looking for information in list form that causes us to say, ‘Hey, that would make an interesting story.’”

Ideas for databases sometimes come from the technically skilled people who, at almost every newspaper contacted for this story, have been hired to fill a new position (or are an existing staffer with a restructured job) as Web developer, data editor or something similar.

PARTNERS IN CRIME

CHICAGOCRIME.ORG offers an idea some newspapers can’t refuse.

The godfather of online crime databases, launched independently in 2005 by journalist Adrian Holovaty, Chicagocrime.org uses public records from the Chicago Police Department and melds it with mapping technology from Google. Users can search the data in numerous ways, including by street, ZIP code and travel routes.

“It’s a good use of public information in a user-friendly manner,” says Chuck Peters, president and chief executive officer of the Gazette Co. in Cedar Rapids, Iowa.

His company’s newspaper, The Gazette, is one of a growing number of papers now mapping local crime

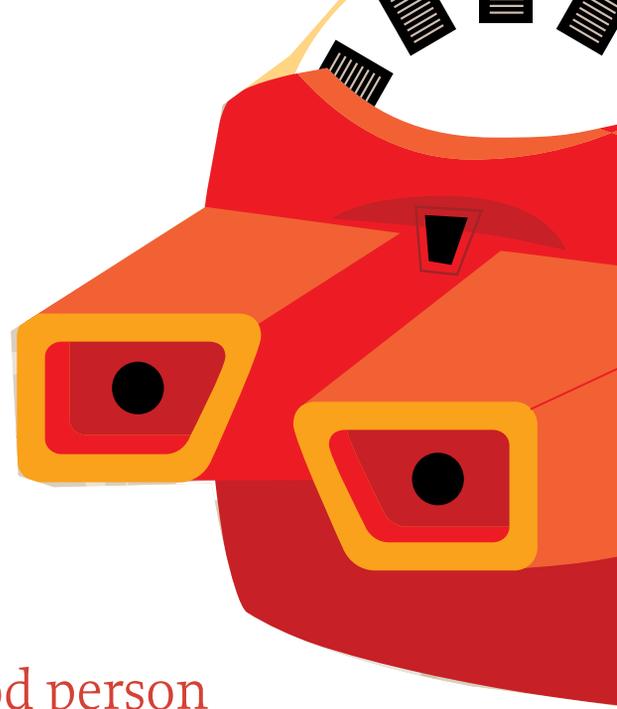
data online (<http://gazetteonline.com/section/crimedata>). The Arizona Republic in Phoenix is another (<http://azcentral.com/CrimeMaps>). Most are updated weekly, but The Fresno Bee ups the ante: Since May 7, the paper’s Crime Map (<http://fresnobee.com/static/crime>) has been refreshed daily.

“The police chief was very cooperative,” says Betsy Lumbye, executive editor of the Bee. It was “easier than I thought to get the information, so the first thing [I would suggest to people] is just ask.” The city police department audits all data and sends the paper daily feeds, Lumbye adds.

At about 10 a.m. each day, the police department exports a comma-separated values file, or CSV file, to

an open Web server. The Bee’s update script (using PHP) then runs at 10:30 a.m., explains Jennifer Ward, the paper’s interactive media editor. The information is saved into a MySQL database. Crime Map, which uses both Yahoo! and Google geocoding services, was developed in-house and “pretty much runs itself,” Ward adds.

Print readers are directed to the Web site through teasers that run near crime briefs and related stories. By June 7, one month after the site’s launch, 30 percent of visitors to the Crime Map page spent three to five minutes there, Ward says, and 20 percent spent five to 10 minutes. On average, Fresnobee.com users spend less than a minute on traditional content pages, she adds.



Doing it Yourself

IN A MOVE few journalists could pull off, Paul D'Ambrosio, investigations editor at the Asbury Park Press in Neptune, N.J., started doubling as webmaster of Data Universe—a staff of one. His passion for data, plus 15 years in computer-assisted reporting and seven years running the Press's intranet, made it a natural transition, he says.

D'Ambrosio creates databases in about two days or less, he says, using the open-source LAMP software stack (Linux operating system, Apache Web server, MySQL database software and PHP scripting language). "It's robust, user-developed, and can handle a tremendous amount of page views per hour—I'd recommend it to any size paper. And the price is right, which is free," he says.

D'Ambrosio learned the system from friends, books, and trial and error, which he wouldn't advocate for a reporter who doesn't already have database experience. D'Ambrosio also builds PHP pages using Adobe Dreamweaver 8.0.

Regardless of the format of original data, it's not too difficult to prepare files for MySQL, he says. He then massages text and checks the integrity of the data. Like many Gannett Co. papers, the Press does a soft launch of a new database midweek, testing bugs before the hard launch and a related print story the following Sunday.

For more information about the Press's Data Universe, go to www.growingaudience.com/resources/article507.html.

“Find a good person or two to speak the database language.”

— TRACI BAUER, DEMOCRAT AND CHRONICLE

Ideas might also come from readers (see story, p. 21). More papers are placing links or forums throughout sites to get feedback; Florida Today has received solid tips with its “blow the whistle” buttons.

And ideas might still start with reporters themselves. At Florida Today, for instance, investigative reporter John McCarthy recently spent three weeks gathering and querying tax data for 2.4 million property parcels from nine counties. The newspaper's systems team created a Web search interface from the table and scenarios McCarthy created, then imported everything to a database (<http://floridatoday.com/taxdivide>; click on the Property Tax Calculator). Like most papers, Florida Today powers databases with PHP and MySQL.

Data arrive in a variety of ways to all papers, including Web scraping

and Excel spreadsheets, CDs and downloads from FTP sites.

Capitalizing on newsroom staff will create even more hearty databases. Traci Bauer, managing editor for multimedia and innovation at the Democrat and Chronicle, says the paper's technology director has built efficiencies into its online calendar of events so reporters and editors can help update it. Gannett's Carroll suggests newspapers start hiring people with degrees in library science, or those who can present databases well visually.

Washingtonpost's Curley often reminds the newspaper industry that the most important part of the newspaper is news, not paper. With that in mind, Bauer offers final advice: “Don't be intimidated by the technology. Find a good person or two to speak the database language, then focus on the data and journalism.” ▣

SOURCES

▲ Mark Bowden, The Gazette, 500 Third St. S.E., Cedar Rapids, Iowa 52401, (319) 398-5869, mark.bowden@gazettecommunications.com

▲ Jennifer Carroll, Gannett Co., 7950 Jones Branch Dr., McLean, Va. 22107, (703) 854-6779, jennifercarroll@gannett.com

▲ Paul D'Ambrosio, Asbury Park Press, 3601 Highway 66, PO Box 1550, Neptune, N.J. 07754, (732) 643-4261, pmd@app.com

▲ Betsy Lumbye, Fresno Bee, 1626 E St., Fresno, Calif. 93786, (559) 441-6207, blumbye@fresnobee.com