

Overwhelming volume of electronic evidence threatens to transform justice system

By JASON TRAHAN

Staff Writer

jtrahan@dallasnews.com

Published: 30 October 2011 11:28 PM

Billions of pages.

That's how much evidence, if printed, the FBI has gathered in a Dallas cyber fraud case. Experts say it is one of the largest amounts of data seized in an American criminal case. Some say it is equivalent to the printed collection of the Library of Congress.

They also say: Expect more of the same.

In an increasingly digital world, law enforcement and prosecutors have more places than ever to hunt for evidence of crimes. Facing a mountain of damaging data, criminal defendants with scant resources to hire their own legal dream team may increasingly be priced out of a trial by jury.

"We're dealing with electronic information in almost every kind of criminal case," said Matt Yarbrough, a former Dallas federal cyber crime prosecutor. "Drug dealers now use QuickBooks. Child pornographers upload to the Internet. Everybody uses email."

Corporate litigants bury one another in hundreds of millions of electronic documents, known as "e-discovery." Private forensic experts can earn \$300 to \$1,000 per hour to slog through it all, resulting in millions in fees.

"If it really costs millions to do that, then you're going to drive out of the litigation system a lot of people who ought to be there," U.S. Supreme Court Justice Stephen Breyer said during a Georgetown law school panel on how electronic evidence is transforming the justice system.

'How we live today'

The Enron investigation and the Dallas-based Holy Land Foundation terrorism financing case are among the high-profile criminal cases that have produced mountains of documents and other digital evidence.

"It reflects how we live today," said Rachel Womack, a former lawyer on the Enron defense team and now a digital evidence consultant in Dallas. "Computer storage is so easy and cheap that people are not quick to throw things away."

But the Dallas cyber fraud case, while complicated, is not a marquee case, except for the evidence haul. FBI agents seized 200 terabytes.

"It's impossible for most lawyers, who don't have large staffs, to deal with huge amounts of evidence," said Tom Mills, attorney for Matthew Simpson, former CEO of a company called Core IP Networks and one of the Dallas defendants.

"It's like having to be prepared to be tested on 3,000 books picked at random — from all the books that have ever been printed," he said. "There is the potential, in the future, for the Department of Justice to be able to make cases that are nearly impossible to defend against. That's because you can't get a grasp on what evidence is relevant far enough in advance to prepare. It can be physically impossible."

In 2010, 19 people were charged in Dallas federal court with stealing \$20 million in long-distance minutes and other telecommunications services. The alleged crimes span seven years.

Fifteen people have already pleaded guilty. Simpson and three others are on trial, which is expected to last into December.

Much of the evidence was seized in 2009 from the 24th floor of downtown Dallas' Univision Tower. An "Internet backbone" facility, it is home to one of the largest banks of commercial computer servers in the country.

Core IP and others were using server space there.

Agents served eight search warrants on homes and businesses, and grabbed about 500 electronic devices, including 280 computers and servers and other gadgets

and storage media, all of which resulted in the 200-terabyte haul.

200 terabytes

"It is patently unreasonable for the U.S. government with a straight face to produce 200 terabytes of data," said attorney and technology expert Ben Kerschberg, who wrote an article in *Forbes* in April criticizing the government for dumping evidence on the defense.

Of the 200 terabytes, agents determined about eight terabytes were relevant to their case. If printed out, that is still the equivalent of 400,000 trees' worth of paper, according to some estimates.

"The volume of data in this matter is beyond anything that I have seen or heard, even in the biggest cases in the nation," defense expert Michael Glazer, president of the Dallas firm Hire Intelligence, wrote in a court filing.

The amount of data is so large that U.S. District Judge Sidney Fitzwater appointed a lawyer, Marlo Cadeddu, just to help wrangle it — a first in North Texas.

Cadeddu took bids from contractors to make the eight terabytes searchable. One firm wanted to charge nearly \$4 million. Another was selected for about \$110,000, some of which was funded by taxpayers.

Evidence rules

By law, prosecutors must share their evidence with defendants and their lawyers so they can prepare a defense.

But evidence rules do not "obligate the government to point the defense to specific documents," prosecutor Candina Heath wrote in a pretrial filing in response to defense protests over lack of user-friendliness and sheer size of the government's evidence pool.

"Although the discovery in this case is voluminous, the government has provided most of the evidence in a searchable or easily accessible electronic form," she wrote.

Not so, said some defense attorneys in the case. In pretrial filings, they accused the government of waiting too long to put the evidence in a searchable database. It took more than six weeks just to convert, court documents show.

Last month, Fitzwater said that the government had been accommodating to the defense. In an order denying a sixth defense request to delay the trial, the judge wrote that prosecutors had made the evidence — "some of which the government was not necessarily obligated to provide" — available to defense attorneys in searchable form since May.

"This is the criminal justice system being dragged into the 21st century," said Jay Ethington, attorney for Casimir Wojciechowski, one of the defendants on trial in Dallas. "We're plowing new ground here."

Richard Anderson, head federal public defender for the Northern District of Texas, said that dealing with electronic evidence is a challenge for his office. And not just in big cases; it costs up to \$50,000 to analyze a hard drive in a standard child porn case, he said.

Often, prosecutors turn over evidence to the defense "in a heap," Anderson said.

"The government takes the position that the way that they organize the data is work product," he said. "We end up duplicating what the government has done to determine what is important and how the case should be defended."

Yarbrough, the former prosecutor, said it pays to work with defense attorneys to help make evidence accessible.

"If you've got the goods, and you can show that early, then you get a plea," he said.

In an era of dwindling government and court-appointed attorney resources, both sides have an interest in working together, Anderson said.

"That's better than them saying, 'There's the haystack. Go find the needle,'" he said.

AT A GLANCE: What's a terabyte?

Technically, it's 1,024 gigabytes, or 1 million megabytes. For comparison, Apple's newest iPhone comes in a 64 gigabyte model. A news reporter might amass about 2.2 gigabytes of email in a year. So, a gigabyte is not necessarily a mind bender.

Neither is a terabyte, from a hardware standpoint. You can buy a 1 terabyte backup hard drive that's about as big as a smart phone for under \$100. It fits comfortably on your desk.

According to Cisco, by 2015, the number of households worldwide generating more than a terabyte per month of Internet traffic will be 6 million. Last year, only a few hundred thousand met that threshold.

Visualizing how much a terabyte of data is, though, is jaw-dropping. For example, in its first 20 years, the Hubble Space Telescope gathered 45 terabytes of data, enough to fill 5,800 DVDs, according to NASA.

If a terabyte of documents were printed out, here are some best guesses for how that would look:

- 50 million pages*

- 20,000 boxes of documents*
- 50,000 trees of paper**
- A 200-mile-high stack of paper***

Source: *Dallas Morning News* research, including * Electronic Discovery Reference Model; **Ralph Losey, a nationally known e-discovery expert; *** Rand Corporation, "The Legal and Economic Implications of Electronic Discovery"