

July 17, 2003

## Picking Up the Pieces

By DOUGLAS HEINGARTNER

**B**ERLIN -- THROUGHOUT the 1980's, Sascha Anderson, a poet, musician and literary impresario, was one of the leading voices to speak out against the East German government and its dreaded secret police, the Stasi.

But his credibility gradually evaporated after the Communist government's collapse as rumors about him acquired the weight of proof: he had been informing on his dissident compatriots all along.

He had been told that his Stasi file had been destroyed. In fact, it was manually reconstructed from some of the millions of shreds of paper that panicked Stasi officials threw into garbage bags during the regime's final days in the fall of 1989.

Now, if all goes as planned by the German government, the remaining contents of those 16,000 bags will also be reconstructed.

Advanced scanning technology makes it possible to reconstruct documents previously thought safe from prying eyes, sometimes even pages that have been ripped into confetti-size pieces. And although a great deal of sensitive information is stored digitally these days, recent corporate scandals have shown that the paper shredder is still very much in use.

"People perceive it as an almost perfect device," said Jack Brassil, a researcher for [Hewlett-Packard](#) who has worked on making shredded documents traceable. If people put a document through a shredder, "they assume that it's fundamentally unrecoverable," he said. "And that's clearly not true."

In its crudest form, the art of reconstructing shredded documents has been around for as long as shredders have. After the takeover of the United States Embassy in Tehran in 1979, Iranian captors laid pieces of documents on the floor, numbered each one and enlisted local carpet weavers to reconstruct them by hand, said Malcolm Byrne of the National Security Archive at George Washington University. "For a culture that's been tying 400 knots per inch for centuries, it wasn't that much of a challenge," he said. The reassembled documents were sold on the streets of Tehran for years.

That episode helped convince the United States government to update its procedures for destroying documents. The expanded battery of techniques now includes pulping, pulverizing and chemically decomposing sensitive data. Yet these more complex methods are not always at hand in an emergency, which is why the vagaries of de-shredding will be of interest to intelligence officials for some time to come.

"It's been an area of interest for a very long time," said William Daly, a former F.B.I. investigator who is a vice president at Control Risks Group, a security consulting firm. "The government is always trying to keep ahead of the curve."

Like computer encryption and hacking, "it's kind of a cat-and-mouse game, keeping one step ahead," he said. "That's why the government is always looking at techniques to help them ensure their documents are destroyed properly."

Modern image-processing technology has made the rebuilding job a lot easier. A Houston-based company,

ChurchStreet Technology, already offers a reconstruction service for documents that have been conventionally strip-shredded into thin segments. The company's founder, Cody Ford, says that reports of document shredding in recent corporate scandals alerted him to a gap in the market. "Within three months of the [Enron](#) collapse at end of 2001, we had a service out to electronically reconstruct strip shreds," he said.

The Stasi archives are a useful reference point for researchers tackling the challenge, though perhaps more for the scale than the sophistication of the shredding. Most of the Stasi papers were torn by hand because the flimsy East German shredding machines collapsed under the workload. The hastily stored bags of ripped paper were quickly discovered and confiscated.

In 1995 the German government commissioned a team in the Bavarian town of Zirndorf to reassemble the torn Stasi files one by one. Yet by 2001, the three dozen archivists had gone through only about 300 bags, so officials began a search for another way to piece together the remaining 33 million pages a bit faster.

Four companies remain candidates for the job, including Fraunhofer IPK of Berlin, part of the Fraunhofer Gesellschaft research institute, which helped develop the MP3 music format. The institute is drafting plans to sort, scan and archive the millions of pages within five years, drawing on expertise in office automation, image processing, biometrics and handwriting analysis as well as sophisticated software.

"It's more than just the algorithms about the puzzles," said Bertram Nickolay, the head of the security and testing technologies department. Indeed, the archive is a massive grab bag of randomly torn documents, many with handwritten and typewritten text on the same page. Combining all these technologies in a project of this scope "is on the borders of what's possible," Mr. Nickolay said.

His system's accuracy rate is about 80 percent. "It will take time for the algorithms to be optimized," Mr. Nickolay said, noting that handwriting analysis began with accuracy levels of around 50 percent, and are now at 90 percent and above.

Some of the companies competing for the job concentrated on the shape, color and perforations of the shreds, while other contenders opted for semantically driven systems, which looked for keywords and likely text matches.

The Fraunhofer plan is to combine its smart scanning software with the know-how of the Zirndorf archivists, who have amassed years of experience working with these tiny pieces of history. After all the shreds have been scanned (at 200 dots per inch), the interactive software will suggest possible matches, which an operator can accept or reject.

While Fraunhofer IPK eventually plans to use a similar technique, several companies say they can do so already.

ChurchStreet's software analyzes the graphical patterns that go to the edge of each piece. First, workers paste the random shreds onto standard sheets of paper, which takes three to seven minutes per page. The pages are scanned, and software analyzes the shreds for possible matches.

Mr. Ford, the company founder, said the ChurchStreet service can recover up to 70 percent of a document's content, although he stressed that the goal was to get blocks of information rather than to re-create the original formatting. The blocks are presented to the client, who determines where they might belong in the overall scheme. "We don't make any guesswork about reconstruction," Mr. Ford said. "We make no assumptions."

ChurchStreet, whose clients are mainly law agencies and private law firms, charges roughly \$2,000 to reconstruct a cubic foot of strip-shreds. A cubic foot of shreds is generally less than 100 pages. Mr. Ford said ChurchStreet would soon offer a service to reconstruct cross-shredded documents - that is, those cut in two directions - for \$8,000 to \$10,000 per cubic foot. A common standard in cross-shredding is particles one thirty-second by seven-sixteenths of an inch, which results in thousands of grain-like shreds per page.

Cross-shredding makes the job a lot trickier, but not for lack of processing power. "The problem is not whether

it's possible with the software, which is possible," said Werner Vögeli, the managing director of the German office of SER Solutions, a company in Dulles, Va., that also competed for the contract to reconstruct the Stasi documents. "The problem is how to scan these documents."

Fred Cohen, a security consultant who reconstructed many pages while working at Sandia National Laboratories, also sees limits. "When you get down to very small shreds, then the numbers start to eat you," he said. "You start to get to where there isn't enough text per shred to be of any use. You've got a completely black shred; whether it's the middle of the cross of a t or the dot of an i, you can't tell."

Adding to the challenge, the smaller the pieces are, the farther apart they can fall, and thus the less likely they are to cluster in a conveniently retrievable form. Security experts also say that using large type (for less text per shred), and feeding the paper into a shredder perpendicular to the direction of the text (so no complete phrases stay together) makes shredding less vulnerable.

Professional document reconstructions are generally recognized by the courts in much the way that fingerprint or handwriting evidence is. An expert may not be able to vouch for the accuracy of the information on a given page, said Mr. Daly, the former F.B.I. investigator, but he can testify that a reconstructed document "was at one time one piece of paper that was cut into little pieces of paper, and now it's back into one piece of paper."

Mr. Daly added that investigators often use reassembled pages as part of a larger forensic puzzle. "Once we have a hard-copy document, we can then go back and look at databases and put in search criteria, and to be able to actually come up with the original electronic version," he said. "One becomes a pathway to the other."

The demand for such investigative services is clear. "I probably get a call every month," said Robert Johnson of the National Association for Information Destruction, an American trade group, from clients looking for "a way to reverse the process."

Other projects, like Mr. Brassil's at Hewlett-Packard, focus on designing a shredder that leaves telltale traces on the documents it destroys, allowing them to be pinpointed later.

In Germany, meanwhile, a decision about whether to proceed with the reconstruction of Stasi documents is not expected before September. Mr. Vögeli of SER Solutions, whose firm withdrew from bidding for the project, said he doubted that financing would materialize. "These documents contain lots of information that might be dangerous to a few politicians who are still active, still in power," he said. "So there's no political majority for any such investment."

Sascha Anderson, the dissident discredited by the files, is among those who hope the project goes forward. "Of course I would have preferred that they weren't found," he said by phone from Frankfurt. "But I realize that it's a unique chance for a society to have access to this information."

And since he was exposed, he said, he has been able to sleep better: "I've ultimately been freed of my burden by history."