
July 15, 2004

STATE OF THE ART

Scanning a Paperless Horizon

By DAVID POGUE

MYTHOLOGY can be hours of fun for the whole family. Remember the one about the guy who flew too close to the sun with his waxen wings? Or the one about the loving couple that Zeus changed into intertwined trees? How about the one about the lady who tried to dry off her little dog by sticking him in the microwave?

But myths don't begin life as myths; in the beginning, people actually believe them. As recently as a few decades ago, for example, you could still hear predictions that computers would one day create a paperless office. Of course, that turned out to be total bunk; the amount of paper consumed in the e-mail age has exploded.

We may never eliminate our file cabinets completely, but it's too soon to give up on that dream altogether. New scanners from Hewlett-Packard, Fujitsu and Xerox/Visioneer can scan stacks of paper, unattended, with marching-band precision - both sides at once, in fact - and convert them into PDF files on your computer, ready for searching, sorting or sending.

(PDF stands for Portable Document Format. It's also known as an Adobe Acrobat format, which you may have encountered in the form of CD-based user manuals or downloadable white papers. Like a photograph, a PDF file maintains the original paper document's fonts, page layout and even pictures, yet you can search for text inside it and copy text out of it. And, of course, you can print it. Maybe Adobe ought to adopt a slogan like "PDF files: The Official Format of the Paperless-Office Myth.")

People in law offices and other paper-heavy enterprises worship the PDF format because it lets them call up a specific page in a specific contract on their hard drive in seconds without having to comb through file cabinets and get paper cuts. Now the same perk is available to anyone.

The Fujitsu ScanSnap FI-5110EOX (about \$400 online, or \$300 after a rebate that's available through September), for example, sits on your desk looking for all the world like a tiny inkjet printer (it's only 5.9 by 11.2 by 5.7 inches). Yet its closest ancestor wasn't a printer but a shredder: you just shove stacks of paper, 50 pages at a time, into the top slot. The machine instantly turns on and begins slurping in pages with astonishing speed. (Fujitsu claims the scanner can process 15 pages per minute, but it lies. You actually get 16 pages per minute.) It spits the pages out the bottom, where, if you're really going for paperless nirvana, you've positioned a wastebasket.

This speed is all the more amazing considering that the Windows software automatically omits blank pages, corrects upside-down ones and even straightens out pages as necessary. The ScanSnap can even scan both sides of the pages simultaneously.

When it's all over, there's a new PDF document - a single icon containing all the scanned pages, in proper order - in your My Pictures folder, bearing a cheerful name like 2004-07-15-17-12-04.pdf. You can change the settings to something that looks a little less like computer programming (like "Scan 001"), but it's too bad the software can't ask you for a title after each scan.

As a handy bonus, you can also feed this scanner a pile of business cards. It ingeniously scans them, translates each card's image into actual text (name, phone number and so on), and sends that contact information into whatever address-book program you use. This feature almost makes it fun to come home from a conference with a pocket full of business cards.

Note, however, that the resulting PDF documents are not, at first, searchable as real PDF documents are, and you can't select text inside them to copy as an excerpt. But Fujitsu provides a full copy of Adobe Acrobat (the expensive program that lets you create and edit PDF files, not just the freebie Acrobat Reader program that everyone has). You can use its Paper Capture command to convert the scanned picture of text into a smart one that can search for the words inside. (The important thing is to find the PDF user's guide on the ScanSnap CD; otherwise, you'll never know how to scan business cards or convert PDF's into searchable ones. Neither the printed Getting Started leaflet nor the scanning software's electronic help even mentions these topics.)

The Xerox DocuMate 252, sold and serviced by Visioneer, is the same idea - a toasterish thing (13.3 by 6 by 12.3 inches) that sits on your desk, always at the ready to accept sheets or 50-page stacks of paper - but twice as expensive (\$865), twice as fast and twice as good.

The speed is absolutely amazing: how does 50 two-sided pages a minute strike you? A digital display on the scanner lets you dial up customizable presets, numbered 1 through 9. Preset 2 might mean "color, double-sided, legal size; send as an e-mail attachment after scanning," and preset 4 can auto-print each scan, effectively turning the machine into a copier. And the documentation is, in a word, spectacular: well illustrated, loaded with tips and troubleshooting and, if you can believe it, all printed up into a nice paperback book.

Every resulting PDF document is searchable and excerptable from the moment it's born - no post-conversion necessary. A bonus program for Windows called PaperPort presents a virtual desktop that collects all of your scans and lets you search your entire PDF collection - a year's worth, say - at once.

Note, however, that both of these compact scanners are designed expressly for scarfing down loose pages. Before scanning, you must first remove staples, clips and all other forms of metal; your documents will feel as if they're going through airport security. You can't scan books or magazines without first tearing out the pages, either. And these document scanners aren't intended to scan photos, although Visioneer says that its machine handles them just fine.

The HP ScanJet 5590 (about \$380), however, has no such limitations. It's much bigger, heavier and uglier - in fact, it looks like a part that fell off an airplane. But good heavens, is it versatile! It includes a glass flatbed (for books and photos), a document feeder (for double-sided scans of loose pages), and even an adapter for transparencies and slides. (Several other companies, including Microtek and Xerox, also sell flatbed scanners with hunchback document feeders grafted onto the lid.)

The advantage here is that you have to buy only one machine for all your scanning needs. The drawback is that this machine is much slower and balkier when it comes to scanning documents, managing only eight pages a minute at best. For double-sided scans, the machine actually runs each paper through the document feeder twice, as though the little elves inside didn't bother to look on the

back the first time through.

Unlike the Fujitsu or the Xerox, the HP model provides its software for both Mac and Windows computers. It can turn your pages into searchable PDF's (or Word files, or e-mail messages) without any conversion step on your part. But for some reason, each PDF document shows up with a dingy gray background instead of white. HP says that there's no fix, at least in the current software version.

HP is another company that has tried to shave costs by eliminating a user manual - and American tech-support workers, at that. (When I couldn't get the document feeder to work at first, my call for help was answered by a someone in India who admirably simulated one side of a "Saturday Night Live" skit. Me: "I can't scan using the document feeder." She: "Can you scan from the flatbed?" Me: "Yes." She: "All right, let's try placing a document on the flatbed." Me: "But I already told you, that works. It's the document feeder that doesn't work." She: "Are you saying that the document feeder does not work?")

All three companies, moreover, have cobbled together a software suite from too many separate pieces - one program for scanning, one for converting to PDF, one for business cards and so on. Not only do you have to install three different pieces of software on your computer (in Xerox's case, on two different CD's), but each program has its own online help, a different software design and a separate entry in your Programs menu. It's a confusing, poorly designed strategy, to say the least.

The machines themselves are a marvel. For the average consumer, you can't beat the speed, value and amazing simplicity of the Fujitsu SnapScan; if you have a small business or you plan to scan daily, the more professional Xerox device is well worth the higher price. Neither machine will fulfill the Vision of the Paperless Office. But maybe it's time to embrace a more realistic vision of the future, one that's actually attainable by scanners like these: the Vision of the Office Without Little Piles of Unfiled Paper on Every Horizontal Surface.

E-mail: Pogue@nytimes.com