

# ACPCUG

**From The  
DealsGuy for  
September, 2002, by Bob  
Click, Greater Orlando  
Computer Users Group**

How much do you concern yourself about buying American to help American jobs? I saw an interesting excerpt from a write-up by a Washington Post writer concerning a typical routine on an average day by a couple he rode with to observe. His theme was US brands owned by European Companies. The couple drove a foreign car with an American name and gassed up at Amoco, owned by British Petroleum. They enjoyed lunch at Burger King, owned by a British Company, Diageo. The husband bought a book published by a french company, Vivendi, and a cassette to play in his car, also produced by Vivendi. In defense, the wife later purchased more gas at a Shell station only to learn it was Dutch owned. She also picked up a bottle of Snapple, made by a British company, Cadbury Schweppes. At home they had coffee, Tasters Choice, made by

*(Continued on page 2)*

## **OmniPage Pro 11 Product Review**

**by Karl Rehak, LVPCUG, email  
karl@idea4u.com**

Optical Character Recognition (OCR) software has been around for many years. Anyone who has had a need to capture an existing document for revision knows the value in having this kind of product to remove much of the burden. A little background of my own experience with OCR software might be in order before beginning the review. I have been a user of this kind of software product for over five years. Starting with the light versions of OCR products with my first scanner followed by two releases of TextBridge. This my second experience with OmniPage Pro. The first OmniPage product I used was release 10.

My general opinion of OCR software is that while it removes most of the effort in capturing printed documents, the OCR process is in itself a very burdensome activity. The software is only capable of doing so much of the effort and any ambiguities in the OCR process requires interpretation that only a human being is capable of providing. As a result I haven't approached an OCR session with delight or as being an activity that would be personally rewarding. Additionally, my OCR sessions are generally infrequent and usually come in bursts that can range from several hours to several days. This means that ease of use is important and the more accurate an OCR program is, the easier will be my brief but intense excursions into their use.

OmniPage Pro release 11 is the first release since Caere was bought and ScanSoft merged their OmniPage OCR program into the TextBridge product. This is also the first new release from ScanSoft in just under two years. Needless to say, and based on my past use of OCR products, I was very interested to see what has been brought to the scene. First and

**September 9, 2002**

**Stark Tech Main Auditorium**

**North Canton, Ohio**

**6:00 PM Password Management  
by Walt Ruthenburg 7:30 PM Q&A**

**8:00 PM Disk Imaging  
by Frank Ramsey**

---

# The Deals Guy

*(Continued from page 1)*

Nestle SA of Switzerland.

There is more, but I found it interesting. With all the mergers and acquisitions, most of us hardly know who we are buying from. Perhaps foreign entities need not engage us in war to take us over, but will simply buy us up while many American companies are moving operations to foreign sites, but I still try to help American jobs. A friend cringes at my "buy-American" attitude telling me he never buys American. He says foreign products are superior to American products because American workers are lazy, conceited, overpaid people always looking for a free handout. Maybe I shouldn't say I was ever an American worker, but I live in America and I am what I am. My American friend also worked for an American company and, I think, enjoys an American pension.

## **Don't Have Acrobat? Then Get Your PDF Files FREE**

When I reformatted my HD, I had misplaced my copy of Acrobat and needed to convert a file to PDF. Sometimes when I get stuck, I call one of my best local resources, which is Hewie Poplock. He told me to go to <http://www.gobcl.com> and get it done there for free. By golly, he was right. **goBCL** is a free online document publishing service by BCL Technologies, the maker of plug-ins for Adobe Acrobat. They can publish your documents into either PDF or HTML.

Not everything is free, but for occasional use, they offer a free service. They also have paid services, but the free one was just what I wanted. I have since found my Acrobat copy, but that service would have been handy.

Hewie also has a Web site where he posts all kinds of information and publishes a newsletter called "Hewie's Favorites," consisting of good URLs he finds. You can get on the list to receive "Hewie's Favorites" each time he sends them out, and it always includes other good information. [<http://www.hewie.net>]

## **Hey, Clean Up That E-mail**

I printed a similar item a few months ago (Stripmail), but some readers have since sent me this information. Since it's their favorite, I'm passing it on. Some people always feel compelled to forward all kinds of stuff. Most, especially newbies, include the >>>, hard

returns and whatever junk there is. These readers recommend using "The Cleaner" on those messages. It gets rid of all those >>> and the unwanted hard returns creating a dreadful looking message. It also offers other features to help.

If you are a "forwarder," get "**eCleaner**" **V.2.01** at [[http://pcworld.com/downloads/file\\_description/0,fid,6492,00.asp](http://pcworld.com/downloads/file_description/0,fid,6492,00.asp)] at no cost and help your fellow humans. There is a later version at [<http://ecleaner.tripod.com>] but this page may have moved by the time you read this. It's only for a PC, but the article's author, Steve Chin, mentions "SpellTools" that is meant for a Mac. He did not say where to find it.

## **Everything Under One Roof, and Cheap Too (for Linux)**

How much did you pay for your Office Suite? Or are you still using an older version because you didn't want to pay that much? If you are using Linux, you could be saving a lot of money with **Hancom Office 2.0.2** just released. I don't know if it will do everything you are doing now, but how much of your present office suite are you REALLY using?

I didn't take the time to check out this new offering from Hancom because I don't have a Linux machine yet, but it offers word processing, spreadsheet, graphics and presentation applications. It has many improvements including foreign language support and is compatible with Microsoft Word format (their words).

It now has Unicode support which has to do with foreign language special characters. They fixed problems affecting Mandrake and SuSE users who had upgraded to KDE 3.0. Fixed bugs when exporting text documents in HTML format.

This suite appears to be only Linux based, but if you own a previous version, your upgrade is free for the download. New users may purchase the software for \$49.95 (CD Case or download edition) or the boxed version with documentation for \$59.95 (regular prices). If you just want to check it out, a free 30-day trial version is available for downloading at [<http://en.hancom.com>]. Some features aren't included in the trial version. I'm not sure I understood the above information on their Web site so take a look. For a dialup connection, this large download may not be very practical, but the trial version is much smaller. See their solution. **Hey, Give Me Back That File, Ya Hear!!**

In Windows 9x, deleted files are normally sent to the recycle bin so that it's easy to get them back if you change your mind. However, not everybody sticks to that method of deleting files. For instance,

*(Continued on page 3)*

---

*(Continued from page 2)*

if you delete a file while holding the shift key down, you can forget about getting it back from the recycle bin. The same is true if you delete it from a Zip Drive.

A short write-up by Ken Marple of Space Coast PC Users Group recommends software called "**Drive Rescue**," A Hard Disk Recovery Tool. The Web site says it will find any lost and deleted data on your drive (e.g. hard disk) even if the partition table is lost or the drive has been quick-formatted. Also recovers lost data that is the result of a system crash. But if the disk has been physically damaged, it cannot recover that data.

Supported file systems are FAT 12/16/32 (used by hard disks, disks, Smartmedia, Compact Flash, Memory Stick and other) and NTFS (incomplete). Supported operating systems are Windows 95, 98, ME, NT, 2000 or XP, and a second hard disk (or network drive) is recommended for recovery. Other amazing capabilities can be seen on their Web site.

The best part is that it's freeware. However, if it helps you, the programmer would like a donation. That's not too much to ask for a good product. The URL is [[http://home.arcor.de/christian\\_grau/rescue/index.html](http://home.arcor.de/christian_grau/rescue/index.html)]. Remember though, if you write to the disk after deleting the data, there is a good chance you would overwrite the deleted data, thus making recovery difficult to impossible.

### **Will The REAL File Manager Please Stand Out!**

Also written up in the same Space Coast PC UG's Journal is a report by Jack Davidson about a great replacement for Windows Explorer. Many people would rather Microsoft had kept the old file manager instead of its Windows Explorer because of the way it works. How would you like to check out a REAL file manager that was originally said to be "a file manager on steroids" when it was created? **Power Desk** is just such a program. Originally, it was a product of Mijenix, but now belongs to OnTrack.

I can't adequately describe the many features of PowerDesk so go to [<http://www.ontrack.com/freesoftware>] and look for it. There are two free programs and two evaluation programs for downloading, including the standard version of Power Desk. If you click on download for the free version of PowerDesk, you will have to sign in and create a password. You may wish to upgrade to the Pro version for \$29.95. I had a problem finding the list of features in the free version. I don't know how many of PowerDesk Pro's features are contained in the free version.

**That's it for this month.** Meet me here again next month if your editor permits. This column is written to make user group members aware of

special offers I have found or arranged, and my comments should not be interpreted to encourage, or discourage, the purchase of products, no matter how enthused, or disgruntled, I might sound. Bob (The Cheapskate) Click [[Bobclick@dealsguy.com](mailto:Bobclick@dealsguy.com)]. Visit my Web site at [<http://www.dealsguy.com>]

---

*(Continued from page 4)*

through the legal words just fine.

The additional pages had slightly less selections to decide, apparently the IntelliScan feature was doing its job. The parts of the documents that contained remnants of holes and slots were passed over by the program, no need to clean them beforehand or eliminate wrong interpretations on the part of OmniPage. Its beginning to get too easy for this program (and me), so for the final scan of the legal document I decided to place it in the scanner at about a 10 degree angle. It scanned fine. The program automatically aligns the text. When I took the document off of the scanner I realized I had also put it in upside down. The program had also fixed that. I became a believer. At this point I was sold, the effect on my overall productivity in the OCR area was great. No more speckle and hole removal or spending countless hours interpreting information. But I figured there had to be one more test I could do.

### **Test Three**

I had achieved my goal of watching the program pass the extreme scanning test, but what else would put it to the test after that? Looking over my bookshelf I saw an old copy of the Bible. The pages are 4 ½ inches wide and 7 ½ inches long, the font was too small for me to measure but is about 7 point or less. Lots of unusual words, numbered verses and italicized footnotes. I placed it on the scanner at a 10-degree angle and hit the scan button. The illustration, "Interpreting the Bible" shows what the program saw. The result was 100% accuracy, 13 suspect words and no changes. On the scale of document quality I put it at about a 6. I'm amazed at this performance and now we're in bonus territory in our extreme scanning test.

### **The Bonus Round**

Added to the program with this release is the ability to OCR Acrobat files. To do this test I found a 'readme' file in my Acrobat folder and scanned it in. It is a 21-page file and all 11,000+ words were input at 100% accuracy with 64 suspect words. Again, the registration symbol and similar characters required confirmation that the

*(Continued on page 6)*

*(Continued from page 1)*

foremost is the claim that the product is more accurate than preceding products and my first question – what does that really mean to me? A feature called IntelliTrain has been enhanced and once again the question – what does that mean to me? Again, because of the intense interaction during OCR sessions, those were the two most significant areas of evaluation to me. Several other new and enhanced features were evaluated as the product was tested and comments regarding their value will be discussed through the review.

The bottom line measure of the value of a quality OCR product is end user productivity. As the quality of OCR products improves and the interpretation of a document's content becomes more and more a function of the computer's ability to interpret information rather than the time spent to have a human being interpret and make decisions, the higher the productivity. That is, in the final analysis, all that matters. So along the way judgments of the OmniPage improvements on user productivity will be highlighted.

### **The Test Begins**

The first test was done primarily to allow me to gain knowledge and confidence in the use of OmniPage 11. I read the instruction manual and reviewer's notes and formulated a test plan to put the product through its paces, but first I had to know what capability was in the software.

For this first test I used the documents out of the OmniPage box. The documents are very legible and of a standard size font, about 12 point Ariel. I rated each test based on the overall quality and legibility of the source document, a 10 is a very easy document and a 1 is an extremely poor quality document. These first test documents were quality 10.

The overall design of the main OmniPage Pro OCR process is built upon the age-old concept of what computing really is. The steps are Input, Process and Output – a simple concept. Add into that each of those processes can be broadly segregated into three different approaches. They can be done manually, automatically or described through a wizard. I tried each method as I tested and found that they are very similar and easy to learn and use. It takes no additional effort to move from one approach to another.

The input describes the source of the OCR, it can be a black and white, grayscale or color document. It can also be a computer file. Introduced with this release is a scanable file type, Acrobat files. More on the PDF scanning ability will be covered later.

The process part of the three elements allows the user to describe what he or she wants as a result. Among the choices are to preserve the page structure of the source, scan it as a single continuous column of information, include or exclude tables or to output to a spreadsheet format. The program also recognizes over 100 different languages. Finally, the output function allows the user to save the information as a file for a word processor or other computer program, output to mail or copy the output to the clipboard. There was another interesting output available – speech. More on that later. The first scans from the manuals went as expected. It interpreted everything with only three suspect words in three pages of scanning. A word with a registration mark (the circle R) next to it is an example of a suspect word. It guessed correctly and I only had to confirm the program's choice. The interpret rating was 100% and again, the document type was a 10.

### **Test Two**

Alright the party's over, time to see just what this program can do. For second test I selected a sampling of a legal document, reproduced many many times, originally bound in a plastic binder with half inch slots, later converted to three hole punched paper. The document because of numerous copies over the years was heavily speckled. I called it the document from hell. This is characteristic of many of the scans I had done in the past and where the personal productivity factor went in the dumper. Lots of interpretation needed and in the older OCR programs the remnants of holes and slots were interpreted as characters. The documents also had to be cleaned of speckles and marks manually with older OCR programs. Further the document had a filing number stamped on it of a different size and slant. Document quality 2 at best on my scale, I decided to name this part of the test 'Extreme Scanning.' The illustration, "Legal Document," shows an image of the document quality. Four pages were scanned and the overall scan accuracy was 93.93%. On the first page it highlighted 17 words. Of the selections it had guessed correctly 15 times and I had only to confirm the choice. I had to select from a provided list of choices twice, and the correct words were in the list.

I'm convinced that lawyers invent words so nobody knows what they are talking about. The net in the OCR test was that the document was loaded with legalese and OmniPage Pro flew

*(Continued on page 3)*

# The Use of Computers in Movies

- Word processors never display a cursor.
- You never have to use the space-bar when typing long sentences.
- All monitors display inch-high letters.
- The most relevant information is displayed in a separate windows right in the middle of the screen, but there's never an Ok button to other way to close it.
- High-tech computers, such as those used by NASA, the CIA, or some such governmental institution, will have easy to understand graphical interfaces. Those that don't, have incredibly powerful text-based command shells that can correctly understand and execute commands typed in plain English.
- *Corollary:* you can gain access to any information you want by simply typing "ACCESS ALL OF THE SECRET FILES" on any keyboard.
- Likewise, you can infect a computer with a destructive virus by simply typing "UPLOAD VIRUS" (see *Fortress*).
- All computers are connected. You can access the information on the villain's desktop computer, even if it's turned off.
- Powerful computers beep whenever you press a key or whenever the screen changes. Some computers also slow down the output on the screen so that it doesn't go faster than you can read.
- The **really** advanced ones also emulate the sound of a dot-matrix printer. (See *The Hunt For Red October* or *Alien*)
- All computer panels have thousands of volts and flash pots just underneath the surface. Malfunctions are indicated by a bright flash, a puff of smoke, a shower of sparks, and an explosion that forces you backwards.
- *Corollary:* sending data to a modem/tape drive/printer faster than expected causes it to explode.
- People typing away on a computer will turn it off without saving the data. (See the opening credits for *The Hunt For Red October*)
- A hacker can get into the most sensitive computer in the world before intermission and guess the secret password in two tries.
- Any PERMISSION DENIED error has an OVERRIDE function (see *Demolition Man* and countless others).
- Complex calculations and loading of huge amounts of data will be accomplished in under three seconds. Movie modems (especially the wireless ones they must be using when they're in the car) usually appear to transmit data at the speed of two gigabytes per second.
- When the power plant/missile-site/ whatever overheats, all the control panels will explode, as will the entire building.
- If a disk has got encrypted files, you are automatically asked for a password when you try to access them.
- No matter what kind of computer disk it is, it'll be readable by any system you put it into. All application software is usable by all computer platforms.
- The more high-tech the equipment, the more buttons it has (*Aliens*). However, everyone must have been highly trained, because none of the buttons are labeled.
- Most computers, no matter how small, are able to produce reality-defying three-dimensional, active animation, photo-realistic graphics, with little or no detailed input from the user. • Laptops, for some strange reason, always seem to have amazing real-time video phone capabilities and the performance of a CRAY Supercomputer.
- Whenever a character looks at a VDU, the image is so bright that it projects itself onto his/her face (see *Alien*, *2001*, *Jurassic Park*).
- Either a *Jacob's Ladder* or a *Van Der Graaf Generator* is absolutely necessary for the operation of new, experimental computers (especially when built by brilliant scientists), although in real life, these devices do absolutely nothing.
- One can issue any complex set of commands in a few keystrokes (see *Star Trek*).
- The internet connects to everything in the movies. You can edit credit records, search hotel registries, lookup police criminal files, search (and edit) drivers license databases, edit social security files and more just using the internet! (see *The Net*)
- Smashing the VDU prevents the whole system from working (see *Speed*).
- You can launch nuclear missiles from any bedroom using an analog modem, but only if you know a single secret password (see *War Games*).

**Akron-Canton PC Users Group**  
**PO Box 2151**  
**Akron Oh 44309-2151**

**Next Meeting - Sept. 9th,**  
**Stark Tech Auditorium**  
**North Canton, OH**

**First Class Mail**

**In This Issue...**

Page 1 – **The DealsGuy** by Bob Click, **OmniPage Pro 11** by Karl Rehak.  
 Page 5 – **The Use of Computers In Movies**

Mailing Label Address Goes Here

**Published by: Akron-Canton PC Users Group**  
**Box 2151, Akron, Oh 44309-2151**

[www.acpcug.org](http://www.acpcug.org)

\* **Membership - 1st year \$25, renewals \$20 per year**

\* **User groups are granted permission to reproduce articles with source credit.**

<http://www.acpcug.org>

Sam will have a serious place in my OCR activity, when the OCR process is done I have a very high level of confidence that it is correct and the voice read back is a redundant activity. Perhaps future releases will find more for Sam to do. The program supports over 100 languages. However, Chinese, Japanese and Hebrew are not among them.

**Conclusion**

All in all ScanSoft has delivered a product matched to its promise. The product is more accurate than previous releases and is easy to use. As each release improves accuracy, the user productivity soars, I would estimate that productivity improvements of an order of magnitude over previous releases can be realized. What used to take 10 hours of effort can now be done in 1 hour. The more complex the document and the poorer the condition the more work the current program will take on, relieving much of the effort of capturing documents.

The program test was a lot of fun to do. The burden of Optical Character Recognition has been lightened. ScanSoft has delivered a significant release and OmniPage Pro release 11 is its name.

*(Continued from page 3)*

interpretation was correct. No changes had to be made.

Added to this release is voice output. I had a lot of fun with this. There are a variety of voice types but the one that was easiest for me to hear was called "Sam." Sam sounds like Steven Hawking. What is the value of this you ask? If you hold the original in your hands, the voice will read what it has and you can do a word-by-word review of captured information. Sam will not have much of a career as a stand up comic, but he can come in handy. Personally, I doubt that

<b>Club Officers</b>			
<b>President</b>	Murray Davis (2003)	440-248-0021	murray5050@yahoo.com
<b>Vice President</b>	Frank Ramsey (2004)	330-699-4711	af Ramsey@yahoo.com
<b>Club Outreach</b>	Rick Lubinski (2002)	330-928-7179	lubinski@bright.net
<b>Secretary</b>	Walt Ruthenburg (2003)	330-283-7867	wruthenburg@usa.net
<b>Program Director</b>			
<b>Treasurer</b>	Joe Krause (2002)	330-484-2338	Joek4646@yahoo.com
<b>Newsletter Editor</b>	Art McRowe	330-666-3189	amcrowe@bright.net
<b>Software Review</b>	Jim Carter (2004)	330-882-4222	jwcarter@bright.net

<b>Future Events</b>	
Sun., Sept. 22nd, 2002	Peter Trapp Computer Show at Tadmor Temple off I-77 to Arlington Road, Akron, OH (Exit 120)