

2/98 FORM Magazine
It's No Optical Illusion

BY JONATHAN ROLLINS

[METROFORM Automates Forms for Real Estate Multiple Listing Service \(MLS\)](#)

A few years ago, Michael Cole and Tom Beede met for their annual brainstorming session. Cole, the owner of Metroform, a distributorship in Rancho Cordova, Calif., and Beede, the president and CEO of MetroList Services Inc., Sacramento, had a long-term business relationship. For many years, Cole had provided forms management and printed products to MetroList, a company that operates a multiple listing service data base for real estate agents and brokers. Agents who belong to the service send information about properties that they want to advertise, and it appears in a data base so other member agents can find matches for their clients who are searching for homes. During the brainstorming session, Cole talked to Beede about high-tech concepts such as bar codes and electronic forms. "Tom's response was, 'That's neat stuff, but it won't do me any good,'" Cole recalls.

What Beede desired was technology that could eliminate the tedious task of manual data entry for the estimated 6,000 real estate agents and brokers who used MetroList's data base. Under the system then in place, licensed agents who were members of the MetroList service completed handwritten profile sheets containing more than 100 questions describing residential properties the agents wanted to advertise. The agent's office then keyed the data into MetroList's host system using computer prompts. "Our whole idea was to transfer the administrative duties from the brokers' offices to us and charge a modest fee," Beede says of his request.

After using the Internet to research solutions providers, Cole returned to Beede with Cardiff Software Inc.'s TELEform package. Today, the OCR and forms processing software automates data entry for MetroList's users. Agents have the option of faxing the handwritten form directly to MetroList's TELEform server, where the OCR software reads the information on the form and posts it to the correct data file. TELEform displays a form and highlights any characters it has trouble recognizing so that a MetroList employee can verify the form and make any necessary corrections. The file is then uploaded into the company's main computer system for listing. With automated data entry, Beede says, the process can be completed in less than five minutes as opposed to 20 minutes when agents key in information manually.

MetroList rolled out its new capabilities to agents and brokers in January 1997. Within 30 days, Beede says, the system was running smoothly, but only after months of testing. "We're the only multiple listing service in the United States that has succeeded in automating a form to the point where it could be faxed in for data entry," Beede says. "Truly, we were scouts on the frontier. It was like being Lewis and Clark. We didn't know what was around the next bend."

The same could be said of Metroform, which was providing its first OCR system. Today, Cole dedicates approximately 90 percent of his time to solutions utilizing OCR. Metroform has successfully evolved from a "typical, but well-diversified business forms distributorship," Cole says, to a provider of high-tech electronic solutions. But, the focus has remained consistent. "The technology stuff is just one more tool in your tool kit," Cole says. "The main objective is still to establish trust with your customer and penetrate the account as much as possible."

[Automating the Manual Form](#)

OCR supplies distributors with enormous potential for penetrating accounts, Cole says. Metroform provided MetroList with TELEform software and a RightFAX fax server, which manages and coordinates the receipt of faxes to the computer system. The distributorship also handled installation, consulting, training and technical support. While high-tech hype is easy to get impressed by, Beede says, Metroform's most invaluable service was still its most basic—testing and designing a form that worked well. In this case, of course, the form had to be both user-friendly and OCR-friendly. "Without a high-quality forms company assisting you with the design," Beede says, "I don't think you can get this to work."

MetroList uses the OCR solution only for its residential property profile sheet, which accounts for 80 percent of total listings received at the company, according to Beede.

When completed, the sheet provides information such as the type of home, the school district the home is located in, site description, type of floor covering, dimensions of rooms and so on. Because the form is so comprehensive, the first considerations for making it OCR-compatible were spacing and font size.

The traditional form used 4-point type and was printed on two extremely tight pages. So the TELEform software could read the redesigned form with acceptable accuracy, Cole enlarged the type to 6 point and added more white space in and between information fields. Extra spacing also was necessary because the redesigned form contained constraint boxes, which help the user write letters that don't touch. As a result, the document can be read more accurately by OCR software.

What resulted from the enlarged font size and additional spacing were four separate double-stub carbonless unit sets glued together at the top. Measuring 8-1/2 x 14 inches, the four unit sets contain three plies each. Agents fill out and fax back the top ply of each set and file the remaining plies. Top plies are made of 20# bond so a pen won't rip through the sheets; remaining plies are 14.5#. To ensure pages from the same listing don't get separated when faxed to MetroList, a consecutive bar code is printed on the first ply of each set. Cole expected objections from agents who were accustomed to completing a 2-page document. Instead, agents welcomed the 4-page form. The larger print and extra white space made the profile sheet easier to read and fill out.

Graphic elements were kept to a minimum, Cole says, because they can affect both the time it takes to fax a form and the accuracy rate when read by OCR equipment. The form was printed in two colors. Red ink signified fields that had to be filled in for the listing to qualify. A cover sheet contains instructions for completing the form so OCR equipment can read it. It asks the agents to print clearly in upper case letters, avoid contact with edges of the constraint boxes and to use a fine or medium-point black pen. The cover sheet also includes instructions for faxing the form to MetroList and provides school district codes, county codes and room dimension codes necessary to complete the form.

Before the MetroList server could process faxed data, Metroform needed to do some programming. Programming included defining information fields as alphabetic, numeric or alphanumeric; defining the maximum number of characters allowed in each field; and determining the field's export path, which essentially gives the data its destination so it posts properly. The form contains a high degree of intelligence. For example, the TELEform program utilizes True Address from the United States Postal Service. This allows the software to compare its reading of a ZIP code against a data base of valid addresses and make changes automatically if it is filled out incorrectly. Or, if Main Street is misspelled as Mane Street, the program uses the data base to check a certain ZIP code for the existence of a Mane Street. Necessary corrections are made automatically.

MetroList integrated its OCR technology quickly and smoothly after the launch, but there were obstacles beforehand. Getting the solution up and running took longer than Cole had hoped for, partly because of Metroform's inexperience and partly because of surrounding circumstances. Three different events were taking place at the same time: a third-party provider was redesigning MetroList's host computer system; a MetroList committee was debating the content of its property profile sheet; and Metroform was trying to implement the TELEform solution and redesign the profile. There were several changes in the profile sheet's

content, resulting in numerous proofs and redesigns before a final decision was reached. "It's like we were juggling three balls at one time," Cole says. "It took the better part of six months. It should have taken six days."

Once the OCR solution was ready, MetroList had problems with its newly-installed Windows NT network. Workstations were addressing the server improperly. Instead of insisting that MetroList have its network provider return to solve the problem, Metroform tried to become network experts overnight, Cole says. "We spent a lot of hours looking foolish with a phone stuck in our ears and talking to four different tech supports," he says with a laugh. Still, Cole calls the hassle a valuable learning experience. "The best way to gain knowledge is to get in trouble," he says. Today, Cole examines his client's network whenever possible before installing OCR systems.

Reducing On-Line Errors

Despite the complexity of the residential profile sheet, the read rate for MetroList's OCR forms is 85 to 90 percent when agents follow guidelines for completion. MetroList also has realized more accurate listings on its data base. With manual data entry, someone in the real estate agents' office types the data into MetroList's computer system. Often, Cole says, the data entry process is rushed so the person can move on to other tasks. Data entry errors show up on-line as a result. With OCR, the only person entering data is a single employee at MetroList who handles verification and correction of the profile sheets. Because the OCR package notifies the verifier of any characters it has trouble distinguishing, mistakes are easy to catch and change. "The data is consistently more accurate," Beede says.

MetroList budgeted a full-time employee to handle verification and correction of the OCR forms, but that hasn't been necessary. After more than a year, the majority of agents continue to enter the profile sheet manually. Only an estimated 10 percent of new listings are faxed directly into the system and read by OCR, Beede says, accounting for roughly 200 profile sheets each month. "It has not proven to be as popular with the agents as we thought it would be," he says. "We thought that this would eliminate all of that manual data entry for them."

Despite the shortfall in anticipated interest, Beede believes in the value of the service. "There's no reason not to stick with it," he says. "We have a good mousetrap, but not many mice. Those that use it love it and can't imagine why anybody would not use it." Many agents who don't use the TELEform option have told Beede they already have employees on staff whose job responsibilities include data entry. MetroList charges \$2 per listing, including forms processing, to agents who use the automated data entry service.

Metroform's involvement with MetroList didn't end with installation of an OCR solution. Cole provides up to 50,000 residential profile sheets to MetroList each year. In the future, he hopes to convert MetroList's other real estate forms to TELEform. Plus, Metroform provides ongoing training and technical support.

Cole believes in the usefulness of OCR and foresees the technology playing a large role in Metroform's future. His other OCR clients include a mortgage bank and a chain of hotels. "If the application requires manual data entry," he says, "then it's a candidate for the OCR solution. The motivation is to automate data entry to the greatest extent possible. The value is immeasurable. Somewhere within the next six months to 16 years, everyone from your 'A' account all the way down to your mom-and-pop accounts are going to be involved with electronic solutions at one level or another."

Jonathan Rollins is assistant editor of FORM Magazine.

Company Capsule

Company: Metroform

Location: Rancho Cordova, Calif.

Founded: 1984

Owner: Michael Cole

Employees: 5

Annual Sales: Estimated \$1 million

Year Company Entered OCR Market: 1995

[Techno Talk: Accuracy Architecture](#)

High recognition rates are the desire of every client utilizing OCR data entry packages. Distributors can largely determine the success of OCR by designing a form that works well with the technology, programming the form intelligently and choosing the correct software.

"A lot of what goes into a good form design for human beings goes into a good design for computer forms as well," says Tim Dubes, director of corporate communications for Cardiff Software^a Inc., a developer of data collection solutions in San Marcos, Calif. Before you design a form for OCR, understand how the form is going to be used, he says, and give it a logical flow, making it simple to fill out. Here are some areas to consider:

Spacing-Provide a generous amount of white space between characters and information fields. If the form is going to be completed by hand, use separation or constraint boxes to encourage neat handwriting and space between characters. "That's going to improve accuracy dramatically," Dubes says. Tightly bunched characters are more likely to be misread by recognition engines.

Font Size-Bigger is usually better when it comes to font size, Dubes says. Most recognition engines can read a small font size, but if the quality of the document is poor by the time it reaches OCR, a small font size is more likely to be misread.

Paper-Variables such as color and size come into play more often with OCR scanners than OCR software, according to Dubes. For example, a scanner may not read a slick paper as well. A solid option is a heavy, pure white stock. Adding background colors to the paper can create "noise" or "static," which will appear as specks on the page.

Non-Read Inks-When a 2-color form printed in green and black is shone under a green bulb of a scanner, the green ink will not be detectable, Dubes says. But if the green ink and bulb vary in shade, static can occur. "Run a test plate, varying shades of a particular color, to see what the optimum shade is for dropping out," he advises. Different scanners use different color filters or bulbs, meaning that non-read inks to be used will vary from scanner to scanner. Using the correct non-read ink can greatly enhance recognition rates, according to Dubes.

Also consider the different capabilities of OCR data entry packages:

Voting Technology-Some processing systems use voting algorithms. These are OCR engines that possess what Dubes calls a "decision management layer." Some recognition systems interpret numeric fields better, while others read alphabetic characters with more accuracy. Using multiple engines allows the OCR system to vote on the proper identification of questionable characters.

Character-Specific Definitions-Software can be programmed to anticipate only certain characters in specific fields. For example, a social security field will contain only numbers, so the software can be programmed to search for the numbers zero through nine. That way, the number "8" won't be mistaken for the letter "B." Some programs also utilize context checking, which helps the system identify a questionable character by looking at surrounding characters and finding the most likely probability.

Dictionary Look-Ups-By associating data bases with certain fields, Dubes says, some software can be programmed to use a field dictionary to find a proper match. For example, a data base might contain the names of 2,000 employees. Characters in a name field may initially be read by OCR as 'Jonas,' but by checking with the data base and discovering there is no employee with that last name, the software determines the real name is Jones.

[Answering the OCR Call](#)

Evolving from a distributorship offering traditional forms to a provider of high-tech solutions can be a nightmare if you're unprepared. "Taking that step is emotional," says Michael Cole, owner of Metroform, a distributorship in Rancho Cordova, Calif. "It's much like making the decision to own your own business." Cole offers the following advice for distributors considering the OCR market:

Training: Personal desire and dedication to learn about OCR is a must, Cole says. "If you and your staff don't buy into the mission," he says, "it's dead in the water." The learning process to become a VAR for automated data entry solutions such as Cardiff Software Inc.'s TELEform products is long and expensive, he says. Cole estimates Metroform spent approximately \$10,000 in training and at least \$50,000 in practice before providing OCR solutions. Partnering with a distributor who is familiar with OCR speeds up the learning process, he says. When Metroform began offering electronic forms, it partnered with a knowledgeable distributor. But when the company offered its first OCR solution, Cole decided Metroform had enough high-tech experience to forge ahead on its own. "Looking back," he says, "I would have partnered with someone. I don't know that it would have saved us hard dollars [in training money], but it would have saved us a lot of time."

Marketing to customers: The selling points of OCR are numerous, Cole says. Automating data entry can save customers time and money. It can eliminate or reduce the number of people dedicated to data entry or free an employee to perform tasks more beneficial to the bottom line. Plus, Cole says, OCR makes data collection painless, meaning companies are more likely to gather in-depth information to help them run their businesses. Some longtime customers may have trouble viewing their "forms distributor" as a technology company. "That's the biggest obstacle to overcome," Cole says. When Metroform sales reps identify a traditional forms customer with an application that fits technology solutions, they usually call in Cole as the technical expert. All Metroform personnel have broad-based knowledge of OCR and other technologies, Cole says, but bringing in one or two people to serve as technology experts aids clients in viewing your company as a high-tech solutions provider.

[Customer Capsule](#)

Company: MetroList Services Inc., a company that operates a multiple listing service for real estate agents and brokers.

Location: Sacramento, Calif.

Founded: 1985

President and CEO: Tom Beede

Employees: 12

Annual Sales: Estimated \$4 million