

# Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

4003 Wood Street ● Erie, PA 16509 ● PH (814) 866-2247 ● FX (800) 507-8981 ● <http://www.documentimagingreport.com>

October 6, 2006

## IBML To Provide Scanners For 2010 Census

**IBML** has been selected as the scanner vendor for the 2010 U.S. Census project, also known as the Decennial Response Integration System (DRIS).

**Lockheed Martin**, the prime contractor for DRIS, plans to use in the neighborhood of 20 ImageTrac IVs to capture images and data from some 90 million handprint-filled forms. 2010 will mark the second time the **U.S. Census Bureau** will rely on automated data collection technology leveraging digital images. It will be the first time those images are captured in color.

"We did very intense testing and studied scanner models from multiple vendors," said Julie Dunlap, Lockheed's program director for DRIS. "We've given each of those vendors a briefing on our decision. The biggest advantage of using the ImageTrac IV is its throughput rate. We actually had to re-architect our software to handle the increased volume being captured by a single unit. In addition, compared to other models, the ImageTracs showed less image distortion."

IBML's ImageTracs are known for their open track transport, which features a flat, conveyer-belt-like, paper-processing path. The ImageTrac IV, which was first introduced at AIIM 2004 [see *DIR* 3/28/06], is the most advanced model. It is advertised as being able to capture 383 8x11-inch

pages per minute. The Census Bureau utilizes 11x25-inch paper for its forms, so the DRIS throughput will be closer to 100 ppm.

"That's almost three times faster than the speed of the Kodak 9520 scanners we used in 2000," said Fred Highland, Lockheed's chief architect for DRIS. "In 2000, we utilized close to 150 Kodak 9520s. For DRIS, accounting for redundancy [in case of downtime], we will probably need 20-22 scanners. This will reduce our overall system maintenance.

"Also, the image quality was significantly better on the ImageTracs than on the competitive models. With large images, like the ones produced from the census forms, you run into some extra distortion issues. Believe it or not, this distortion is exaggerated when you capture images in color. The ImageTrac does a pretty good job of minimizing this distortion compared to the other models we tested."

### **Advantages of electronic dropout**

The use of color images will enable Lockheed to implement its internally developed electronic color dropout technology in DRIS. For the 2000 Census, a colored bulb was used to drop out the background color. The images were then captured in black-and-white.

According to Highland, there are two main advantages to electronic dropout. "The first is that it enables us to drop out multiple colors at once," he told *DIR*. "This will enable the Census Bureau to use multiple colors on its forms to improve their visual appeal. For the dress rehearsal [planned for April 2008], the Census Bureau only has plans to use a single-colored form, but the option is there to add more colors to the final form if they wish.

"In addition, electronic dropout is more sensitive to different shades of color. This means that if someone uses a red pen to fill in a form with a red background, their answers won't be removed with the background, as they would be on a scanner with a red bulb. This should reduce the number of exceptions that have to be processed manually."

### **Improved capture software**

In an environment that receives some two million forms on a peak day, reducing exceptions is a major goal. For the 2000 Census, some 88% of the form fields were able to be captured through OCR, ICR, and OMR processes, without any manual keying. This is a percentage that Lockheed seems pretty satisfied with. As a result, it has not made too many changes to its core data capture system. Once again, technology from **Océ ODT** will be leveraged for the OCR/ICR portion of the system.

One change that is being made is that, because color images are being used, the TIFF, Group 4 file format used for the 2000 Census will be replaced. "We are still planning on capturing images at 200 dpi, because that resolution represents the best balance between our requirements for image quality [*many OCR/ICR engines are tuned to work best at that resolution*] and restrictions on image size," said Highland. "We are considering using a JPEG format. We have looked at JPEG 2000, but we like the universality of regular JPEG. It's typically easier to go with a widely accepted standard."

Another change is that OMR technology from **Optimum Solutions, Inc.**, which was used in 2000, will not be used in DRIS. "That decision is strictly based on budget constraints," said Highland. "We were going to implement a voting scheme combining Lockheed's own OMR technology with theirs, but we've scaled that back and are only using our OMR."

A third change is the implementation of new key-from-image technology for exception documents. In the 2000 Census, Lockheed had utilized **Captiva** keying technology. "For DRIS, we're using keying technology we first developed for our Canadian Census project in 2006," said Highland. "For the 2000 Census, often, if a form was damaged, it couldn't be registered properly. As a result, we couldn't use the image. The form had to be kicked out and data had to be keyed from the paper. This was an expensive exception process."

"With DRIS, if we can't register the form, we can just send an image of the full page to a keyer who can work off that. Of course, it helps that we're using color images, because they're better to view for keyers than black-and-white ones."

### **Some details being ironed out**

According to Dunlap, a final decision will be made in late 2007 as to exactly how many scanners will be purchased and at how many sites they will be deployed. It also has to be decided whether the scanners will be leased, as they were in 2000, or purchased outright. For the 2000 Census, service bureaus were set up in Pomona, CA, Baltimore County, MD; and Phoenix, AZ, in addition to the Census Bureau's permanent facility in Jeffersonville, IN.

Originally, it was thought that an Internet forms option may affect paper volumes, but according to Highland, online forms will not be part of the dress rehearsal. He could not comment on the Census Bureau's ultimate plans for e-forms.

Another planned technological upgrade that

appears to be in jeopardy, at least if you believe what you read in the Washington-area press, is the Field Data Collection Automation (FDCA) system, for which **Harris Corporation** has been awarded a \$600 million contract. The contract calls for development of a mobile computing system for collecting data from households that do not submit paper forms.

"We are working very closely with Harris," said Highland. "We are designing DRIS to accept data from mobile computers. We also have to communicate with the FDCA system to let it know who has responded and who has not. Similar information has to be communicated to the printing contractor for the follow-up mailings that precede the door-to-door efforts."

### **Experience paying off**

Respondent information is contained in a bar code on the returned envelopes. These bar codes are scanned as soon as possible after arriving at the processing center. "In 2000, we also slit open the envelopes automatically," said Highland. "We haven't made a hardware decision in that area for DRIS yet. Regardless, after the bar code data is captured, we have about a two-month window to scan and capture the data from the forms."

Once all the form data is captured, Lockheed will save the color images for the Census Bureau, which apparently still hasn't worked out a long-term archiving agreement with **NARA**. If you remember, after the 2000 Census, there was a disagreement about who would be in charge of maintaining a digital archive. The result of this disagreement, after some grandstanding in Congress, was that the digital images were copied onto microfilm, before being transferred to NARA [see *DIR 12/21/01*]. For Lockheed's DRIS contract, we reported that archiving was included as an "unpriced option [see *DIR 10/21/05*]." According to Dunlap, "it remains an option and has not even been discussed yet."

Dunlap concluded that the DRIS project is running more smoothly than the project for the 2000 Census due to the increased experience of everyone involved. "2000 was Lockheed's initial census project involving digital imaging," she said. "Since then, we've had the same team work on the U.K. [2001] and Canadian [2006] census projects. In addition, this isn't the first time around for our customer [the U.S. Census Bureau] with digital imaging."

For more information:

<http://www.lockheedmartin.com/wms/findPage.do?dsp=fcc&ci=13413&rsbc=0&fti=126&ti=0&sc=400>  
<http://www.census.gov/procur/www/2010dris/index.html>