IT-DIRECTOR.COM

When 'scanner' becomes an inadequate product category

Published: 19th December 2008

It is easy to assume that a scanner is a scanner is a scanner. The sole objective is to be able to scan a document of some kind—an image or a typescript for example—and capture the contents to computer for subsequent use. The only decent arguments to be had about the subject are document sizes handled and the imaging quality possible. In some specialist areas of the scanning world however, as DRS Data Services Limited has found, the truth is far removed from this picture.

Take the world of elections as a prime example. Getting elections right, especially Government elections, is of crucial importance. One only has to think of some now infamous 'hanging chad' to understand the potential for argument and dispute over a result, and the potential consequences that can follow. One option in solving this problem is the traditional one—armies of people counting votes by hand, with further armies overseeing the process to ensure fair play and accuracy. The alternative is to use a mixture of the oldest and latest scanning technologies, with software tools that can provide as much as possible of the necessary fair play and accuracy, while being seen to work transparently by and with the human authorities overseeing the process.

Such requirements demand a mix of hardware and software technologies that are in an entirely different league to the humble scanner, in much the same way that a Boeing 747 and a wheelbarrow are both forms of transport. The markets where such capabilities are important are certainly niche, but they all share a common requirement—the need for real-time data capture to the highest levels of accuracy, at the highest throughput of scanned items, all completed to the highest levels of auditable transparency in operation.

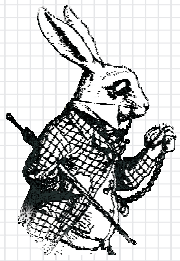
In fact, it is almost a disservice to call the DRS PS960 a scanner at all, though the company does so for the sake of convenience. In most respects it is an all-encompassing, holistic, data-capture environment where the scanner is one, albeit major, component. The needs of its specialist marketplace demand that a wide range of technologies are employed together, such as Optical Mark Reading (OMR), Optical Character Recognition (OCR), Intelligent Character Recognition (ICR), Barcode Reading and yes, Image Scanning as most users know and understand it. The reason the unit has such a multiplicity of data capture options is that many of its applications require multiple input formats, not just to get a result, but to validate that result and resolve any subsequent issues or disputes that might arise over the user's intentions.

Using a election ballot paper as a example, the complexities of the voting process and the need for operational transparency in operation demonstrate why the PS960 has been designed as it has. In fact, the process starts with the paper itself. Ideally, this has to be of a definable quality so that it can withstand being passed through a series of drive rollers without distortion. This is important because the ballot must carry a set of code markings that identify it as a legitimate ballot. Those without the correct markings are routed to the scanner's Bad Hopper for human adjudicators to examine as are ballots printed on poor quality paper. As well as being prone to distortion and tearing, poor quality paper can also be the cause of multiple pick up, where more than one ballot is fed through into the reading system. This problem normally has two possible effects in many scanner systems—it either causes a miss-read if not detected, or it jams the scanner. The DRS solution to the problem has been a software routine that returns a multiple pick up to the loading tray once detected, where the process is retried two more times. In the vast majority of cases, this clears the problem without causing a jam.

As part of the holistic data capture environment, DRS has its own specialist printing facility for ballot papers and similar customer data forms.

The printed check marks are recognised by the OMR system, which, at the same time, also reads the handwritten marks made by the voter if that voting system is being used. Software in the PS960 can identify the orientation of the ballot paper from the printed check marks, removing reading problems, or incorrect reads from papers read upside down from prescribed normality. The OMR system can also be used to read barcodes that identify the ballot papers in some way—for example the specific election where the vote is being cast .

OMR provides a very simple reading system, but only up to a point. As voting options grow, OMR forms can become more complex and difficult for voters to understand. Often, as with a



IT-DIRECTOR.COM

When 'scanner' becomes an inadequate product category

Published: 19th December 2008

preference voting system such as the Single Transferable Vote (STV), it is inappropriate. This leads to the need for ICR, for the requirement is to read the voter's preferences as denoted by ranked number. The system can be set to identify just the areas of the ballot paper where the voter gives that preference and, again, the characters are read on the basis of the paper's orientation, removing possible errors.

Image scanning is also used in the PS960 but in specific ways, for using it for whole page scanning in high-speed data capture systems has several weaknesses. It is, for example, an over-kill in terms of data capacity used in the process. It can also be slower in operation. But it is often the case these days that some image scanning is needed, for example where a fingerprint or visual image of the voter is considered as an option in positively identifying individuals. Here the PS960 can be programmed to scan only the element of the ballot paper where an image will be found.

However, it should also be possible to selectively scan the whole paper should it be necessary. This capability has been built into the PS960 for the purpose of dispute resolution. If, for example, a ballot paper has OMR markings that are indistinct or slightly mis-placed or if an STV voting number is undecipherable by the system, or perhaps beyond an accepted range, the system records an image of the whole paper, and sends the image for manual adjudication. This speeds the process of resolving disputes because the scanned image can then be forwarded via the network to all the necessary adjudicators representing both the authorities and the individual parties to the election. They can all view the image at the same time, rather than sequentially with a single ballot paper, speeding the resolution process.

The mention of speed raises an important issue with the PS960, for DRS is happy to acknowledge that it is not the fastest scanner in terms of the time taken for any form to travel through the machine. Speed, however, is not the same as throughput, and in the world of real-time data capture the primary issue for users is to get access to accurate results from the data capture process as quickly as possible. In the case of an election, therefore, the real test is the time between voters casting their votes and a validated, accepted result being declared. This is based on achieving a high overall throughput of ballot papers rather than the brute speed of the scanner. This comes from the flexibility of the scanner itself, coupled with such factors as the ability to clear many paper jams by re-trying, the availability of software that can accommodate a wide range of different form types and data formats at the same time in a single pass, and the ability to manage problem resolution online and on-screen.

Size and cost also become an important issue, particularly in the world of elections, for the PS960 has already found a strong market in countries where elections are still relatively new and the national infrastructure is less well-established than in Western Europe or the USA. Here, the DRS target has been to produce a table-top system that has a similar power consumption to a typical PC (the fundamentals of which lie at the heart of the PS960) and can be delivered at a cost-effective price. Coupled to the high-throughput of the machine, DRS feels it has hit on the right formula not only for election result management, but many other applications areas where that combination of throughput, flexibility and scanning accuracy in real time data capture need to be applied to get the right result such as with a census or examination.

As one final part in that all-encompassing, holistic data capture environment DRS targets is the company's bureau service. Here, it can take on the complete management and provision of the data capture service a user requires, from designing and printing the forms through to delivery of the final data.

Martin Banks,
Practice Leader - Data Centre & Mainframe,
Bloor Research