



White Paper

The Undeniable Importance of Paper Capture and SharePoint®

Russ Edelman, President & CEO, Corridor Consulting

**Nuance Document Imaging Group
2009**

Executive Summary

The book titled, “Business@the Speed of Thought”, published in 2000 by Bill Gates, predicted that the Paperless office would ultimately triumph over the paper office. While progress has been made in this capacity, the reality is that a far distance remains in achieving this bold vision. In part, the physicality of paper still holds a special place for many as the complete transition to digitization remains uncomfortable.

Yet, progress has and continues to be made towards digitization and over the past few years, two complementary and unique forces have allowed critical mass to develop in this new digital world. The first force is driven by the unquestionable business need to convert paper documents into their digital equivalents. The second is the need to methodically (and sometimes not so methodically) store these “digital assets” into a centralized repository. On the second count, it has become abundantly clear that a new leader has emerged and that leader is Microsoft® SharePoint® 2007.

The power of a rich and robust repository, coupled with an intuitive and capable digitization platform will allow organizations to achieve new heights in terms of productivity and efficiency; a state that moves us all closer to Bill Gates’ digital vision.

The Force of Digitizing Paper

As organizations around the globe look to achieve improved efficiencies, many recognize that the digitization of paper rates high on the priority scale. While there are a number of “creature comforts” associated with paper; the reality is that it is inherently inefficient. It is time-consuming to physically look through file cabinets, boxes and folders and to make copies for distribution. Distributing documents for review and collaboration means expensive postal costs, overnight carrier costs and inefficient and often insecure fax transmissions. The editing of paper documents is also not effective as they must be retyped and paper form completion takes a lot of additional time. Finally, paper is prone to disasters, both big and small, as it is only redundant when copied. And for most, this is not an exercise that any of us prefer to pursue.

For all of these reasons, the trend has become abundantly clear. First, wherever possible, avoid paper and second, if paper is inherent in a business process, it should be digitized as quickly and efficiently as possible. To accommodate these needs, users are now scanning papers to their desktops (on-demand), using Multi-Function Printers/Devices (MFPs) for centralized scanning and in rarer circumstances, relying upon production grade scanning stations.

For those unacquainted with the “nuances” of digitizing or scanning paper, it is important to point out that going through the actual scanning process represents only half the battle. The other half is introduced when users attempt to convert the scanned “picture” of the document into a meaningful and rich format. Specifically, scanned documents come to life and present an entirely new level of value if the document

PricewaterhouseCoopers Research Study

- There are over 4 trillion paper documents in the U.S. with an estimated growing at a rate of 22% per year
- Companies spend \$20 in labor to file a document, \$120 in labor to find a misfiled document, and \$220 in labor to reproduce a lost document
- Of all the pages that get handled each day in the average office, 90% are merely shuffled
- The average document gets copied 19 times
- 7.5 percent of all documents get lost, 3 more percent get misfiled

image can be converted to searchable and formative text. To accomplish this, the right tools are needed to enliven a scanned picture of the document into a meaningful document asset.

The Force of SharePoint 2007

With a collection of meaningful document assets, the importance of organizing them and making them available to a larger audience becomes increasingly important in order to leverage their value. This was and still is being addressed in divergent fashions across and within organizations. Through the use of a variety of Enterprise Content Management (ECM) tools documents and information can be made available to everyone. Since the introduction of these ECM Tools, industry pundits, vendors and users have predicted their mass proliferation and wide spread adoption. Yet despite, these predictions, ECM tools typically appear in small pockets throughout organizations.

“SharePoint is one of Microsoft’s fastest growing products”

*Thomas Rizzo
SharePoint Director
April 2008*

And then SharePoint 2007 was released and the ECM industry changed in ways that were truly unexpected.

Prior to SharePoint 2007 arriving, its predecessors (SharePoint 2001 and SharePoint 2003) met with mediocre results at best and followed the adoption patterns of most ECM products. With SharePoint 2007, the newly introduced functional capabilities, coupled with a simplified user interface, opened a floodgate of demand that to this day, is still not satisfied.

SharePoint 2007 has for the first time, ushered us into an era of unprecedented and near ubiquitous ECM adoption. Globally, organizations are for the first time facing the reality of corporate memories that extends deep into the realm of unstructured information such as document assets; assets that often serve as the lifeblood of an organization.

Key Challenges with Paper, PDF and SharePoint 2007

Organizations create large amounts of information in various formats including paper, PDF files and MS Office document types. SharePoint 2007 shines most brightly with MS Office documents as organizations can retain this business critical information. Many organizations must also comply with industry regulations that specify how information is stored, archived, and disposed of once the documents no longer have business value or legal relevance. SharePoint 2007 includes these business critical records management features and yet, with all of SharePoint 2007’s innate capabilities, it is not perfect. In this regard, SharePoint 2007 does not include the necessary tools that allow scanned documents to be digitized and optimally managed. While virtually any file can be stored in SharePoint, including traditional graphical files (TIFF, JPG, BMP, etc.) and PDF files, integration with scanners and advanced scanning engines are only provided through third party tools. Additionally, a cumbersome and technical process must be undertaken to enable PDF files to be searchable in SharePoint.

With the introduction of third party tools, users can scan documents into SharePoint and save them, typically as TIFF, JPG or PDF files. This approach is typically the simplest as no conversion takes place between the scanned image and meaningful text and metadata. To raise the bar in this capacity means

that when documents are scanned, a more sophisticated and intelligent process takes place which extracts the information in ways that allow for searching, and workflow automation.

The challenge of extracting meaningful information from documents does not actually stop with the process of digitizing paper into images. It actually extends to PDF files as well as it is often equally as difficult to extract information from PDF files and place the information into alternative formats such as Word. In most cases, the formatting of documents is lost when utilizing tools to extract information from PDFs to their editable equivalents in MS Office formats. In the absence of being able to do so, productivity can be reduced as users must go through the laborious process of retyping and reformatting the information in MS Office format.

In the optimal scenario, users would be able to easily move information and the associated formatting from physical, digitized and PDF documents into editable files such as MS Office and to do so through simplified conversion, collaboration and reordering of the files.

The Power of OmniPage OCR Technologies

Nuance Communications, Inc., a Microsoft® Gold Partner, is the world leader in OCR, document conversion and PDF technologies. Nuance® OmniPage® has long been the market leader in speed and accuracy for creating rich editable documents from scanned paper. Nuance, in cooperation with Microsoft was also a pioneer in PDF to Microsoft Office conversion and again leads the way in document conversion to native Office Open XML formats for Office 2007. Now OmniPage 17 has been streamlined to improve the collaboration process and to automate paper and PDF related workflow processes and to do so by being SharePoint 2007 aware. Listed below is a summary of key OmniPage capabilities that ease the transition into the digitized world:

Easily Capture Documents with Networked Multifunction Printers

- Get ready to scan quickly and accurately with any network MFP - OmniPage can route documents to your desktop or SharePoint 2007
- OmniPage increases the value of documents scanned with networked MFPs by converting documents directly to Microsoft Office formats

Easily Create and Assemble PDF Documents

- Create Searchable PDF documents so you can search for text inside your scanned documents from your desktop PC or from inside of SharePoint 2007
- Communicate in color with 8x smaller scanned PDF Files
- Create fill-able forms from scanned forms or static PDF forms
- Create secure PDF documents by applying password protection to scanned documents

Communicate Clearly and Precisely

- Redact (black-out) or highlight important information
- Utilize Nuance's built-in RealSpeak engine to convert paper and PDF files into WAV audio files that can be listed to anytime and anywhere

OmniPage and SharePoint 2007

OmniPage is the cornerstone of advanced technologies developed by Nuance Communications, Inc. Nuance has had a unique and in-depth relationship with Microsoft products. Through collaboration with Microsoft, Nuance invented the PDF conversion category in 2003 and the innovation continues. Nuance imaging products now deliver Office Open XML and XPS formats for Office 2007. And working with SharePoint 2007 is no exception. By utilizing OmniPage organizations can move closer towards the vision painted by Bill Gates.

Users are now able to leverage the capabilities of OmniPage and transform digital files directly into SharePoint 2007. Once in SharePoint 2007, users will be able to easily find the digitized file in one or more formats. Leveraging SharePoint 2007's security platform, OmniPage users will only be able to see what is available and most relevant to them. SharePoint 2007 workflows can be initiated to reflect important business processes to further enhance productivity.

About the Author – Russ Edelman, President & CEO, Corridor Consulting

Russ is the founder, President & CEO of Corridor Consulting, a systems integration and consulting firm focused specifically on Microsoft SharePoint 2007 and the Enterprise Content Management (ECM) industry. Founded in 1997, Corridor has designed and implemented over 300 ECM solutions and its clients include organizations such as First Data Corporation, Gulf Oil, Anheuser-Busch, Nike, Microsoft and many other great companies.

Russ also remains very active as an industry analyst where he observes and helps to define key industry trends. He contributes to periodicals including InformationWeek, CIO Magazine, AIIM Infonomics/E-DOC and KMWorld. He is a featured speaker at industry conferences and sits on a number of emerging technology committees in the ECM industry. These include AIIM's Emerging Technology Group (EmTAG) committee; a group established to provide feedback and direction to the AIIM board and the international AIIM constituency. He performs a similar function for ARMA's Technical Advisory Committee (TAC) group.