



KYOS Systems Inc.  
20 Mall Road, Suite 200  
Burlington, MA 01803  
781-272-1401  
[www.kyos.net](http://www.kyos.net)  
[contact@kyos.net](mailto:contact@kyos.net)

## **The 5 S's required for a Successful Document Data Management Solution**

Or, How to become a hero and not a goat...because many scanning and archiving projects fail to provide the intended value for the company. Make sure your efforts don't waste time and money, only to have a document repository that fails to support your end users.

In this paper we discuss the 5 S's necessary for a successful solution: Strategy, Scanning, Storage, Search, and Sharing.

### **Legacy Thinking**

Most companies and CIO's focus on just two (Scan and Store) of the five "S's" and therefore are missing a chance to provide long lasting value from any paper reduction project. By thinking through the 5 S's from the point of view of DDM (Document Data Management), i.e., getting and using data from documents (see DDM white paper), the optimal value proposition of both saving money and achieving robust access to data and scanned images is possible.

Most companies start with the second S, Scanning. This is a mistake. While Scanning (the 2<sup>nd</sup> S) may seem simple and straightforward (which it is as a core activity), it can rapidly evolve into unwanted complexity. Unwanted complexity means increased cost and labor as various creative workarounds are devised to compensate for discovered deficiencies, e.g., how much indexing is required to make search effective for everybody, how much staffing to achieve desired cycle times, dealing with missing pages, establishing what constitutes a complete file, etc.

Storage – the other general piece to 'run of the mill' scanning projects – is critical, yet is often an afterthought of the project. Either an existing legacy system is "available", a general purpose content management system (more suitable for electronically generated documents) is purchased, or a database is modified. In these cases, storage is not optimized to support Search, a major component for system success. A key issue is that Scanning is normally the province of Record Managers and Storage of IT managers. It is critical that a solution encompasses the input and optimization of all players in your organization.

Therefore no matter where you are in your scanning 'life cycle;' contemplating scanning for the first time, in the middle of an implementation, or have been scanning for awhile, KYOS always recommends taking a step back and analyzing the total System (perhaps a 6<sup>th</sup> S) and devising the Strategy to attain all that you can from your document imaging and archiving activities.

## **Future Thinking**

### **Strategy**

The Strategy for attacking a scanning project begins with asking the question, "What is the maximized value you can expect to gain from your imaging and archiving?" Analysis of this kind leads us to examine a number of aspects of your total system, both IT and human, and importantly, where they intersect in workflows. For example:

#### **Approach with the end user in mind:**

1. Who and what roles in how many workflows within your organization require access to the images?
2. Why will they need access to the images? How often?
3. What level of access do you need to provide? Are there protected or restricted fields of information that need to be managed?
4. What is the access and search process? How much time are you allocating for each search, access, and data usage? Web service, internal servers, etc.
5. How much time in training in system usage are you allocating?

#### **Assess the cost issues:**

6. How much data entry, i.e., transferring data from image to another application is done with the images? Is this done at scan time – or later in a workflow..?
7. Are there targeted fields within each form or document that could benefit from automated data conversion? And how do we get those into the system?
8. How much business intelligence are you currently extracting from your imaging? What is your cost to get to that data?
9. Can you expect your storage and management solution to support data mining, auditing, and analysis? What is the cost of such a system?
10. Do you know the true cost of your imaging and archiving activities, including all FTEs that interact with the system?

#### **Are you "Future Proofing":**

11. Where do you see your system 5, 10 years from now?
12. Do you have a scalability plan?
13. How will you ensure search and usability functionality as your archive grows?
14. What is your life cycle management policy?
15. Is your current solution flexible enough to support workflows as they evolve?

This analysis may seem overwhelming, and the pros at KYOS will help. Before we start any project, we determine the end users needs and requirements. The above are just some of the questions to which KYOS delivers answers. Once a map of the entire scanning-to-storage-to-search-to-sharing process is made, you are ready to optimize your scanning solution.

### **Scanning (or Image Processing)**

Scanning involves image processing; either at scan time or once in the archive. And processing is critical to making the documents and the data within the documents useful to the organization.

Scanning itself comes in two flavors each with two variations. Scan can be outsourced to a qualified vendor, or performed internally. KYOS can walk you through breakeven analysis which will show you when to outsource and when to retain internally. Configurations include either scaling for larger volume backfile (digitizing historical records) conversion or day forward conversion. A final choice involves whether to centralize scanning or have it be distributed in multiple workstation sites. This latter choice has remarkable impact on local workflows and staffing. A clear understanding of what is to be achieved depends significantly on timeframe demands, e.g., does the document data need to be accessed right away e.g., in hours or days, or even weeks or beyond? How much processing is required for what workflows, and what is the key process bottleneck?

### **Storage**

This is truly the Achilles heel in any Scan-Store-Search-Share operation. Most organizations attempt to use content management systems that are ill-suited to their downstream document data needs. Document management systems primarily used to manage electronically generated content rarely work well in the effective storage and efficient search of imaged documents. The failure of these systems comes from their predication on the 'one document-one PDF static object premise.

Page metatagging (typically Date and Name) is the primary mechanism by which to find a particular page. Yet metatagging each page is labor intensive and expensive, and so many organizations opt for minimal metatagging. Furthermore – if done at scan time, the scanning process can be seriously slowed down, leading to increased cycle times and document inaccessibility. If done post scanning, then separate workflows and additional FTEs are required.

This creates a significant flaw and 'data penalty' that threatens the success of your system. That is, the more you store, the more difficult it is to find what you need on demand. Limited metatagging is a severe constraint on the scalability and utility of your system. To compensate, the Search component of the system becomes increasingly expensive and ineffective. Unless you have a crystal ball and can predict future search and workflow requirements, a limited metatag approach, while perhaps saving some cost on the front end, severely limits system value and potential. By only having a couple of tags, the richness of search is seriously curtailed.

So look for a solution that automatically tags each and every element on every form page, creating a multi-dimensional metatag universe by which each page can be found through any of its data elements and most uniquely, each data element itself can be found and used independent of the form itself, allowing incredibly powerful data manipulation and usage. This Storage strategy creates dynamic lightweight data objects that individually have meaning, and in aggregate, define the form from which they derive, and furthermore, when combined with data from other forms and documents create new levels of meaning and intelligence, and thus value, from your archive.

### **Search**

Most organizations are actually drowning in data-data that can be productively transformed into useful information, if they just knew how to find it.

Other than providing better distributed access, standard digital archiving solutions make absolutely no improvement over paper based search. Users must open folders, open thumbnails, and visually verify that they first have the right document, the right page, and the right data. They must then do something with that data, either mentally process, data enter into another application, or aggregate with information from other sources. There is absolutely no difference between performing this with PDFs and paper. No wonder many users feel that they are more efficient using paper. Interviews with end users such as doctors, lawyers, and other non-IT professionals indicate that many believe that they are actually more efficient searching through paper than PDFs!

Because of the dominant paradigm of static object (paper) to static object (PDF) used by many vendors, Search is an overlooked area for optimization. Inefficient search is the single largest stumbling block to organizational acceptance and efficiency. Inefficient search leads to many of the unwanted complexities that lead to altered workflows both upstream and downstream of scanning.

Look for solutions that are “future-proof” in that they store and utilize all available information on the pages. In addition, make sure that the user interface is intuitive (easy to learn) and yet powerful in its capabilities. The optimal interfaces allow Users to perform and configure rapid and complex queries simply by pointing and clicking, removing the need for users to learn and memorize complex lexicons and data dictionaries in order to find and use data

### **Sharing**

Document data dependent workflows proliferate due to increasing organizational informational and regulatory compliance needs. More people need access to the data contained within the pages. How do you manage your data to provide robust yet secure data that optimizes all the different workflows and different workers with different needs, backgrounds, and expertise-now and in the future? Again, standard solutions default to having each person having to visually search, validate, and extract data, leaving the rest of the operation to the end user. This leads to wide variance in performance, productivity, and satisfaction or frustration.

The key to effective sharing is having the flexibility to selectively flow and control document data to feed and power all the diverse workflows within your organization. How can you do this with static objects? What is required is the ability to capture specific data elements and flexibly append business and processing rules as needed to support what you need to accomplish with that data. What you need is access is not just access to the document, but data within the document; being able to aggregate data across forms, through time, as needed to flow into informational workflows.

### **Summary**

- **Strategy:** Have you considered your overall System needs; from workgroups processing the work to the various end users who have need for the document data? Have you considered what the system needs to do 5 years from now?
- **Scanning:** What is the cost of ownership of any system-outsourced or internal? What is the mix of automation and manual effort required to achieve targeted turnaround times and scalability?
- **Storage:** Does your solution truly support Search for all end users? Can your solution allow auditing, aggregation, and data mining?

- **Search:** Do you know what all your potential end users need to find and use the system for? Have you made it easy for people to find and use document data?
- **Sharing:** Have you created workflow maps that determine level of access and effort to find and use data? What is your document and document data security policy?

Whether you are just considering or already implementing a scanning solution, it is never too late to step back and ask these and other questions of your system and your vendors. Especially when you consider the cost, the FTE commitment, the inevitable workarounds, and the future needs of your organization.

Once you have mapped and catalogued your answers to these questions, you are ready to begin optimizing or building an optimized system that will more than likely create a Successful Solution for you and your organization.

### **The KYOS Solution**

Getting the 5 S's correct means that another S: Solution, will be a Success for you and your organization. KYOS combines its proprietary and award winning platform solution, KYOS TransFORM™ along with its deep operations expertise to help you save money and create greater value. The combination of our expertise, industry leading technology, workflow engines, and experienced staff can help you reduce your manual processes by as much as 50%.

KYOS offers industry leading search of form and document data. Because we 'atomize' each page of every form by extracting and tagging each element on the page, we revolutionize document data search and manipulation. KYOS builds intuitive computer-human interfaces that require almost zero curve learning. Users perform rapid and complex queries simply by pointing and clicking, removing the need for users to learn and memorize complex lexicons and data dictionaries in order to find and use data. We help you control the variance by mapping the data needs of each workflow and configuring the Search and Share components to maximize productivity.

Whether you are just beginning to scan or have been scanning for awhile, we can help you design, build, and extract greater value from your paper assets and PDF/image libraries.