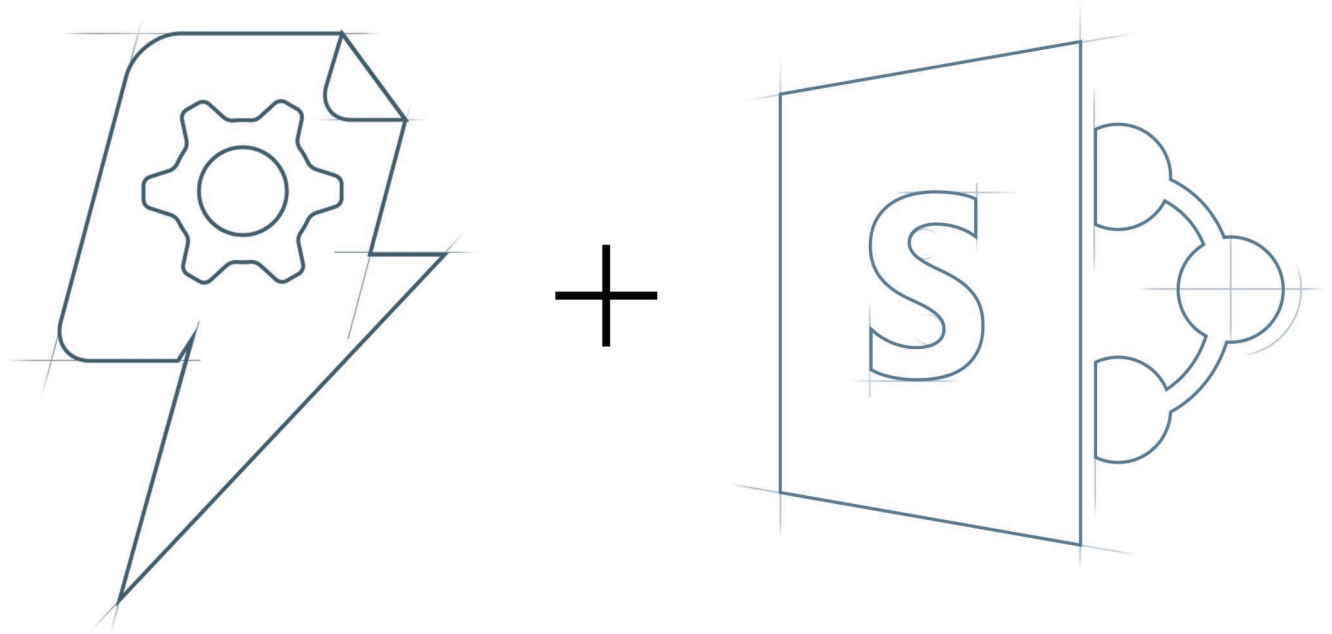




PSIGEN software **inc.**
the **science** of document capture

[White Paper]



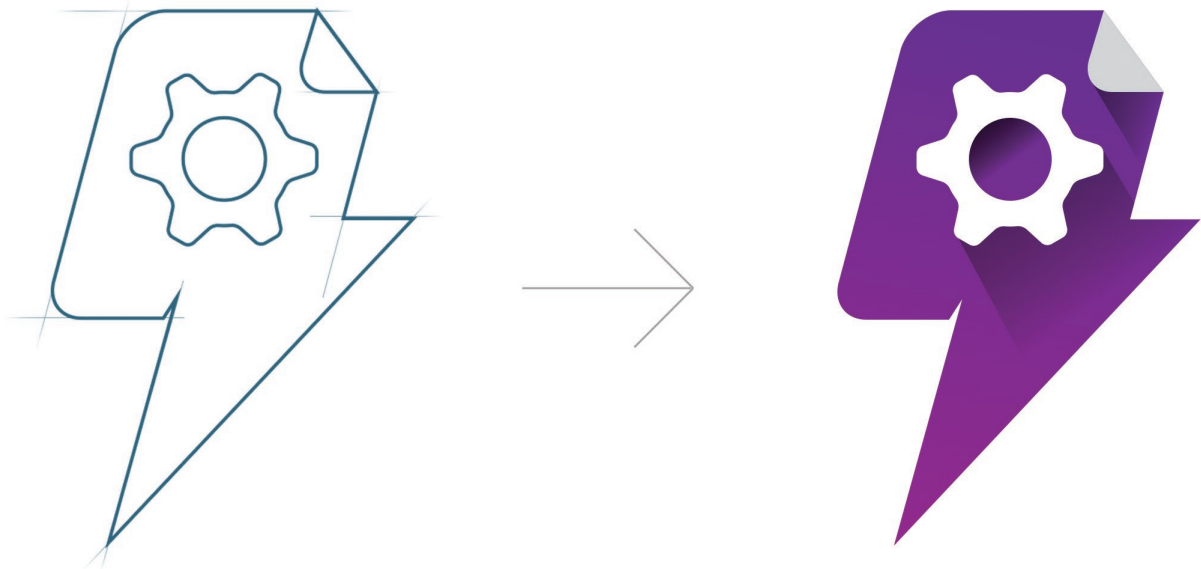
Leverage SharePoint with PSICapture

A White Paper Courtesy of PSIGEN Software, Inc.



Table of Contents

Intro to PSICapture	1
Intro to PSICapture for SharePoint	1
Key SharePoint Features	2
Key PSICapture Features	3
Integration with your SharePoint Site	7
Example Solutions	11
Bringing it All Together.	13



A White Paper Courtesy of PSIGEN Software, Inc.



Intro to PSIsapture

PSIsapture is PSIGEN's flagship product and leads the industry as the premier advanced document capture software. PSIsapture helps companies automate document processing by extracting data from ingested pages and utilizing it to name, index, format and route the document along the electronic path to storage. The lasting result is a more productive end user, using fewer mouse clicks and keystrokes, while accomplishing complex and powerful document workflows.



Intro to PSIsapture for SharePoint

Microsoft SharePoint has served as a premier content management tool for the past 15 years. In that time it has flourished and expanded, becoming a staple for many organizations. SharePoint is more than just a document repository. It allows for employee collaboration, task management, easy and customizable search, and much more.

Although SharePoint is an extremely useful business tool, many organizations simply aren't using it to their full advantage. According to a recent study conducted by AIIM entitled "Connecting and Optimizing SharePoint — Important Strategy Choices," 26% of respondents admit their SharePoint project has stalled and 37% say their SharePoint project has not met their original expectations. With the success SharePoint has experienced and the number of businesses that continue to deploy it, these numbers seem strange. Why do so many organizations struggle to optimize their SharePoint experience?

An essential workflow component that can be easily overlooked in the deployment of a SharePoint project might be the key to understanding some of these issues. SharePoint has a plethora of functions, but it can only go so far as a standalone application. If your business is importing documents to SharePoint, whether in the form of scanning physical papers or importing digital files, there's danger in jumping in headfirst without a plan and without a program that can effectively help you make the most of SharePoint.

PSIGEN's PSIsapture advanced document capture software provides the ultimate solution for capturing documents and their contents and routing them into SharePoint libraries. PSIsapture boasts an expansive assortment of cutting edge features, making it an unmatched SharePoint scanning onramp. With PSIsapture, your business can truly unlock the potential of SharePoint as a content management system. You'll no longer have to worry about wasting time searching for the files you need, and your organization will be able to make use of the platform to collaborate on projects and share important information. This white paper outlines how SharePoint's unique and highly customizable organizational and search structure can be integrated into an advanced capture workflow through PSIsapture, providing users with a seamless end-to-end solution—from scanning documents to accessing, editing and collaborating on them in SharePoint.



Key SharePoint Features

The easiest way to understand PSIsapture's integration with SharePoint is to review the various features of SharePoint that will ultimately be used to integrate the two programs and facilitate the capture and indexing of documents to SharePoint. The following section is a basic overview of the main SharePoint functions that will be used to simplify the capture and migration process.



Managed Metadata & Term Stores

Determining the terms that will define your information is a fundamental feature of SharePoint that ultimately makes SharePoint sites and libraries manageable, searchable and functional. Managed Metadata is the overarching organized set of description words that are planned and set throughout the enterprise. Organizing SharePoint using Managed Metadata makes the SharePoint experience much more manageable and customizable based on what you are looking for. Term Stores, which consist of a hierarchical organization of terms—Groups, Term Sets, and individual Terms—are defined by the Managed Metadata service.

Using metadata to tag content allows SharePoint users to effectively use the search function to filter based on the specific applicable metadata. SharePoint also allows users to dynamically create views based on metadata fields, a function that can be very useful to individual users, as long as the Managed Metadata has been populated correctly and consistently throughout the SharePoint site.

Managed Metadata is an important part of your configuration between PSIsapture and SharePoint. As you will read later, PSIsapture has the ability to dynamically set index fields based on mapping to your SharePoint site's Managed Metadata fields. By setting up these fields in SharePoint, you'll make it possible to greatly reduce the amount of time spent entering data during document capture, and set your organization up for a fast and accurate automated capture process.

Content Types

SharePoint Content Types provide a way to classify documents in your SharePoint site by type. Whereas Managed Metadata tags files with terms to aid in organization and search, Content Types provide a way to indicate what type of document you are dealing with. Be it an invoice, contract, memo or other form, setting up Content Types for your SharePoint site is another way to help manage the information stored there. During configuration, PSIsapture automatically displays the Content Types available in your SharePoint environment.



Document Sets

SharePoint users often lament the difficulty of using SharePoint's folders effectively. Document Sets can help organize your files in your SharePoint site and are an effective alternative to folders. Document sets allow you to manage multiple documents as a single item, which can be useful for groups of documents such as employee records, multiple documents surrounding a specific case, and the like. Document sets can be assigned metadata as well—a feature that makes them superior to folders in most cases.

Folders

While it is true that many SharePoint users discourage the use of folders for organization in SharePoint, it isn't a universal rule by any means. There are certain instances when you might want to use folders within your SharePoint libraries. Sometimes folders are used as enhanced security for documents that should only be accessed by authorized users, as they allow you to easily set permissions. They are also useful for storing very large amounts of content. Since a single List or Library in SharePoint only supports 5,000 items, folders can come into play with databases containing very large numbers of documents.

Key PSIcapture Features

PSIcapture has over 40,000 features, making it an extremely robust software product. Fortunately, utilizing PSIcapture does not require extensive knowledge of all of these features. In fact, PSIcapture is quite an easy platform to understand and very simple to use, once you are familiar with its main components. The following are the main features that will be used when you integrate with SharePoint.

Importing to PSIcapture



Capture
Module

PSIcapture is the ultimate onramp software for scanning and capturing documents destined for SharePoint, providing granular control of the scanning workflow from start to finish. One of PSIcapture's many great qualities is its ability to import your documents in almost any way you can think of. Scan them from any dedicated scanner or network scanning device, or import them from network folders or SharePoint WebDAV. Wherever your documents are located, PSIcapture has a way to get them scanned in or imported to begin the workflow.

PSIcapture also provides an auto-import function, which monitors specified folders for new activity and runs newly-detected documents through the PSIcapture workflow automatically. This function is especially useful if your organization has multiple scanning locations, or employees who scan documents but are not involved with the capture process.



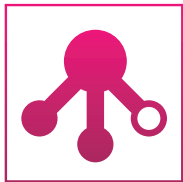
Making Sense of Your Documents: OCR



OCR
Module

The powerful engine that facilitates the conversion of scanned images into meaningful text is Optical Character Recognition (OCR). What some people fail to realize when scanning their documents into SharePoint is that the process of scanning itself doesn't create anything more than an image. Thus, directly scanning and entering your business's important documents into SharePoint doesn't create searchable PDF documents. There is no way for the content of these images to be understood by SharePoint because they do not contain meaningful text. In order to convert a scanned document to understandable text, you need to employ OCR technology. PSICapture's OCR functionality is highly advanced and accurate in converting scanned images to searchable PDFs—a function that is crucial to your overall workflow.

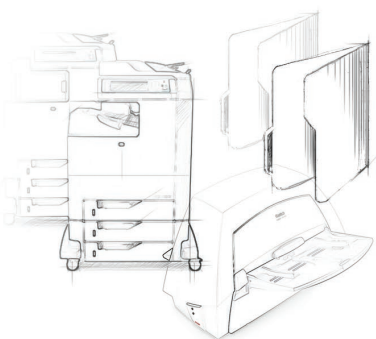
Classification with ACE (Accelerated Classification Engine)



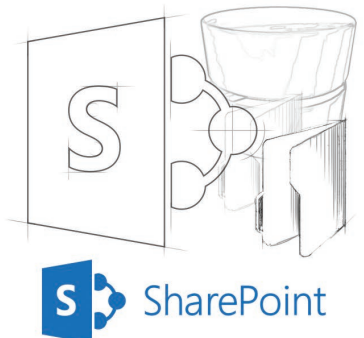
Classification
Module

PSICapture's Classification Module provides intelligent classification of forms and documents. By parsing the document for a particular word or phrase, the software is able to determine what type of document is being processed. Within the Classification Module is a powerful automation feature, the Accelerated Classification Engine (ACE), which allows users to classify new capture profiles in the middle of a batch, without having to interrupt the workflow. ACE can also connect to the user's database to preconfigure Classification forms and rules using already existing data.

By classifying the document discipline, you can then set rules within PSICapture for required fields. This metadata can be used within SharePoint for populating Content Types, Term Stores, Columns, Document Sets and Document Names. By reducing the configuration time and empowering the business users to leverage ACE, the capture profile within PSICapture will grow as the users continue to operate the system. This in turn will increase user adoption and SharePoint consumption.



PSI**capture**





Indexing and Data Extraction



Index
Module

Arguably the most painstaking and time-consuming part of the document capture process is populating index fields. Left as a completely manual process, you might have employees spending hours every day inputting data into index fields, but without indexing you merely create a digital landfill. Of course, the data collected within this step is what makes your final product—the PDFs in your SharePoint site—searchable, making it a crucial step in your workflow. This is the data that will be mapped to your SharePoint organization structures—Managed Metadata, Content Types, etc. Fortunately, PSICapture has a full arsenal of tools that can turn this nightmarish task into a simple, pain-free process.

There are several different options for automating the indexing process, including manual input, point-and-click entry, Zone OCR and barcode population, and database lookups. Note that if your SharePoint site contains information that can be used to populate index fields, you will want to use the database lookup option to help speed up and automate the indexing process. This process is outlined in the next section, “Integration with Your SharePoint Site.” If your SharePoint site does not currently contain relevant data that would assist in automating the indexing process, the options outlined in this section are recommended.

Advanced Data Extraction (ADE)



ADE
Module

The ADE function is an extremely useful tool that provides the most sophisticated method of auto-indexing and routing of documents to SharePoint. ADE pre-OCR's documents and runs a pattern match across the entirety of the document set. With ADE, it doesn't matter where the information is on the physical document—the whole page gets parsed for patterns. ADE will even parse multiple pages for information. This is a very efficient method of indexing, especially for documents where the information needed for indexing is not in a consistent place on every page. It's also useful if you are looking to collect data for business intelligence purposes, such as parsing multiple files for a specific word, phrase or number.

ADE is extremely accurate and is a terrific tool to take advantage of if you want to truly automate your workflow with very few initial setup steps. ADE is an advanced module, rather than a core PSICapture module, that must be purchased in addition to PSICapture.



Routing Sheets

Routing cover sheets provide another method for collecting data for indexing. Cover sheets allow the end user to designate the desired folder and naming structure of the document being processed via barcodes and check boxes. It requires a minimal amount of upfront end-user intervention in order to facilitate automated extraction.

The routing sheet is placed on top of the document set you are scanning, then the documents are scanned and the automated indexing process begins. PSIs capture will auto-extract the data specified by the check boxes and/or barcodes. This information can include the SharePoint Site and Library, the Content Type, folder structure and more. The data specified on the routing sheets can be existing sites and folders, or they can be dynamically created by PSIs capture.

OCR Assisted Indexing

OCR can extract text from a document in a variety of ways. You can run OCR for an entire page, or run it for one or more defined areas, called zones, to increase processing efficiency. You can upload a document template and then configure zones on the template to determine the location on the document from which the information will be extracted. You can choose to run OCR on certain pages of a document or on the entire document. OCR assisted indexing has a variety of customization options, including auto correction settings, character filtering, and using Regular Expressions to further define the type of information you want indexed.

Migration



Migration
Module

Once the PSIs capture workflow has completed its task of importing, arranging and indexing your documents, Migration publishes your documents to your SharePoint site. The migration configuration is highly customizable depending on your specific needs. This step is explored in-depth in the next section. <Dynamic Migration/ Index Metadata> <Publish metadata into LOB applications such as ERP by publishing into database tables or by creating properly configured load files for ingestion.>



Integration with Your SharePoint Site

The PSIsapture configuration with SharePoint is robust, user-friendly and capable of making even the toughest capture and indexing tasks fully automated.

PSIsapture's integration with your SharePoint site will map your indexed data to your SharePoint Managed Metadata, Libraries and Content Types, so that when your documents are converted to PDFs in your SharePoint site, the indexed data will be attached to the appropriate fields designated in SharePoint.

PSIsapture Capture Profile

PSIsapture has a pre-configured Capture Profile template for documents that will be migrated into SharePoint. When you import your documents into PSIsapture, you can select the desired SharePoint template from the list of options and PSIsapture will configure the workflow for a SharePoint migration.

SharePoint Versions & Authentication Types

Connecting your SharePoint site to PSIsapture is easy. Simply enter the site address, username and password, and choose the version of SharePoint your organization uses. PSIsapture supports SharePoint 2007, 2010, 2013, 2016 & Office 365 Online. After PSIsapture has established a connection with your SharePoint site, it will load a list of your existing Libraries. You will have the option to choose to use these Libraries for your workflow or to create new ones, which will appear in your SharePoint site once you have completed the entire workflow and migrated your documents.

PSIsapture offers several options for authentication. PSIsapture connects through NTLM web services by default, but the user can also use the client side API to connect. You can also choose to authenticate with ADFS, which allows the authorization of users from an existing Active Directory. ADFS is a secure claims-based authentication process that allows quick authorization without the need to create and memorize usernames and passwords. Another option is Forms-Based Authentication, which is supported when user credentials are stored and managed in a provider such as SQL Server. PSIsapture also has a Mixed Mode authentication option, which allows for multiple ways of connecting. You might use this option if you need to set up different sets of users who will require different methods of authentication.

Advanced Connection Settings

Advanced authentication

☒ Enable advanced authentication

Claims authentication type: ADFS

ADFS identity provider configuration

Identity provider URL:

Trust version: Use WsTrust Version 2005

SharePoint online authentication

☐ Use ADFS identity provider

Ok Cancel



Lookups & Indexing

If you want to use existing data from your SharePoint Site to populate index fields during the indexing step in PSIsapture, the lookup configuration, located in the Advanced Indexing step, is the tool for the job. You can choose a list lookup, which returns all the records from the selected row within a table or view, allowing you to select the record you wish to use to populate the selected index fields. You can also choose to do a query lookup, which uses values in specified

index fields to run a query in your SharePoint database table or view, returning only records matching those values. PSIsapture also has the ability to access specific Term Stores from your SharePoint database for the lookup. You will specify the Term Store ID and Term Set ID to find all of the terms in the Term Set. You also have the option to fill in the Parent Term ID, if you want to narrow your search down further.

Mapping Indexed Information

The data that has been extracted during the indexing step can be mapped to your SharePoint Site's existing Metadata columns. Any field that has been indexed by PSIsapture can be mapped to your SharePoint Site's fields. Once this is done, captured documents that are published to your SharePoint site will be tagged with your Managed Metadata, making them easily searchable.

Migration

In the migration step, you determine the specifics of your document exportation to SharePoint. These can include the Document Library, Folder Content Type and Document Content Type. You will also determine the image output type, naming structure and any folder design you might want to utilize.

Field Type	Field Name	Key/Return	Mapped Field Name
Document	TaxID		
Document	Status		

Options to configure the Microsoft SharePoint Direct migration

SharePoint system options

Site version: SharePoint Online (Office 365)

Site address: https://psigen.sharepoint.com/sites/test/Training

User name: jessica.rivers@psigen.com

Domain: psigen.com

Password: *****

Web: /

Migration type: Migrate to Document Library

Document Library: Existing

Folder content type: Existing

Document Content Type: Existing

Image file options

☒ Migrate image files

Image source: Main Image

Image output type: Multi-Page Tiff Image

JPEG quality: 75

Tag to output:

All page behavior: One Document with all tags

OCR and non-image file options

☐ Migrate OCR or non-image files

Files to migrate: All OCR files for Document

OCR workflow step position: 1



Naming Structures

Choosing a method for naming your documents is a crucial step in your document capture workflow. Because SharePoint's dynamic search functionality is such a key element of the program, it's important to set up a file-naming scheme that will be easy to identify when your employees do a search in the SharePoint site. PSICapture allows you to use various constants and values from your documents to populate your document name field. Thus, you could use a constant such as Invoice and a variable like the exact value contained in the Invoice Number index field to provide the naming structure for your files.

Example:

Constant: Invoice

Value: Invoice Number

File Name: Invoice_3021

Microsoft SharePoint Migration Configuration

Settings for custom naming of outputted files

Custom file naming options

Field Type	Field Name	File Name	Type	Field Name/Constant	Mask	Field Type
Document	Customer	Add Field	1	Const	Invoice	None
Document	Invoice Number	Add Constant	2	Field	Invoice Number	Document

Buttons: Edit, Remove, Move Up, Move Down

Current pattern: Invoice<Invoice Number>

Buttons: Save, Cancel

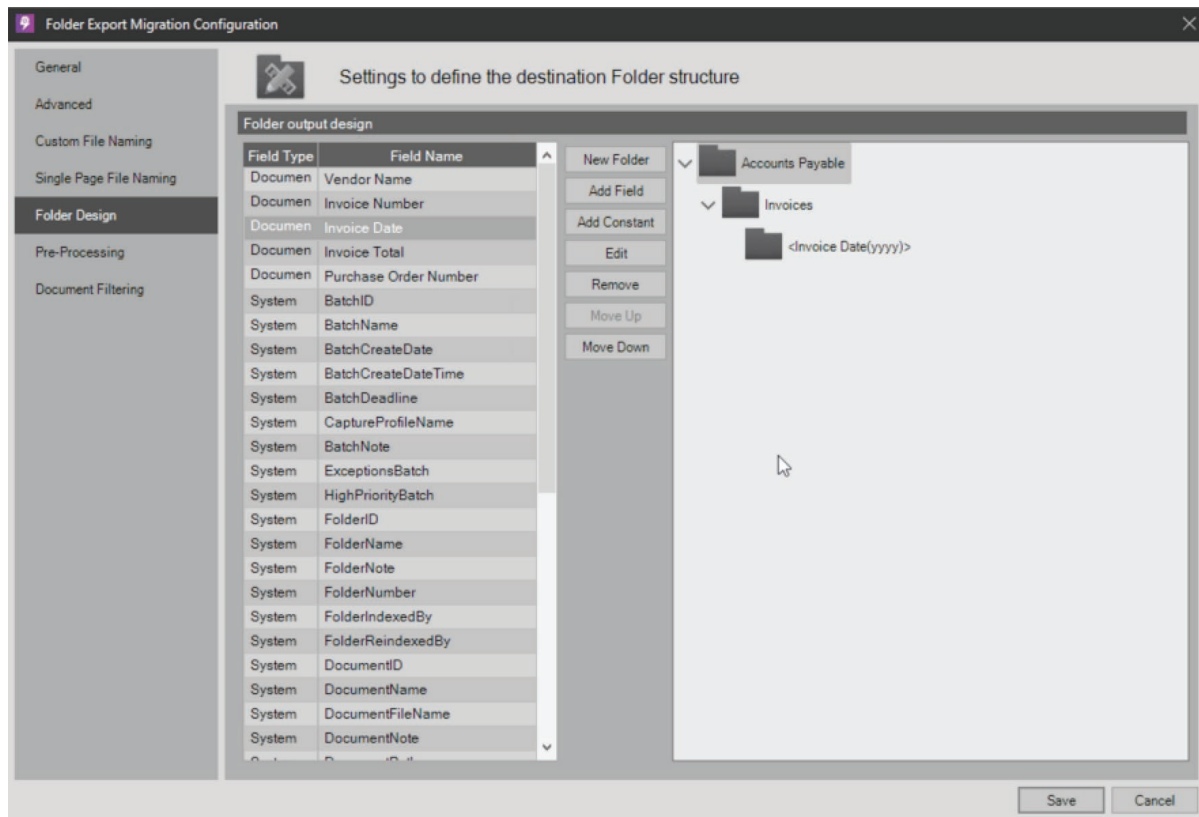


Folder Design

PSIcapture gives you the ability to create folder structures for your documents within your document workflow, without ever having to create or arrange them in your SharePoint site. You can decide exactly how you want your documents to appear once they are migrated into SharePoint using the Folder Design tab in the Migration step of your workflow. Remember that you don't have to have pre-existing folders in SharePoint in order to arrange them in PSIcapture. If they don't already exist, PSIcapture will create them for you and you will automatically have access to them when you open up your SharePoint site.

Example:





- Accounts Payable
 - Invoices
 - 2014
 - 2015
 - 2016
 - Expenses
 - 2014
 - 2015
 - 2016





Running the Workflow

Once you are satisfied that your fields have all been configured to your liking, complete the workflow by confirming your configuration in the migration step. This step completes your PSIs capture configuration, and PSIs capture will begin running through the workflow, stopping for any manual functions you specified so that the user can verify accuracy and perform the manual functions where needed. Finally, the completed documents are published to your SharePoint Site, where they are immediately available for viewing.

Workflow		
1	Capture / Import	
2	Index	
3	Quality Assurance	
4	Migration	

Example Solutions

PSIs capture's SharePoint integration isn't just a good concept in theory—it's a solution that has worked exceptionally well for many customers. These real-world success stories prove that the PSIs capture/SharePoint solution is one that can make a profound impact on your business's efficiency, productivity and bottom line.

Microsoft Legal

Problem: Microsoft's Legal and Corporate Affairs (LCA) division had a paper problem that needed a resolution. The department was buried in paper records, and strict requirements for their record-keeping process created major business bottlenecks.

Solution: All incoming files are scanned into PSIs capture

Solution: The department devised a scanning process to back-scan the existing paper files with PSIs capture. Once the documents had been scanned, PSIs capture used OCR to index the documents, then Classification to organize each document into one of six grouped document sets. The documents were then migrated to SharePoint Records Center as searchable files. PSIs capture was also linked to the department's database, which allowed the software to extract information to auto populate names, birthdates and other information, which was uploaded as metadata to the Records Center. After processing the 12,000 current active files, they also moved Microsoft Legal's existing electronic files to SharePoint. Now when the legal department receives files, they are simply scanned into PSIs capture from a scanner, or imported from an electronic format, where they are classified and indexed, then migrated seamlessly to Records Center, where they can be easily accessed through a simple search.



Silicon Valley Bank

Problem: Silicon Valley Bank (SVB) had accumulated loan and deposit documents for over 4,000 clients, all of which were in paper form and could only be accessed by employees at the company's headquarters. With 47 bank branches requesting documents from headquarters, SVB was receiving up to 100 requests a day, each taking up around 15 minutes of employee time.

Solution: Loan and deposit documents are scanned in, given a standardized name and published directly to SharePoint. PSICapture also creates document libraries and folders within SharePoint to improve searchability. Employees at SVB's branches worldwide are now able to access files instantly from the SharePoint site. The elimination of document retrieval has increased productivity for SVB's 1,200 employees.

Haulfryn Group

Problem: Haulfryn Group was dealing with an old scanning system that was slow, low quality and was unable to effectively store their documents. They needed a solution that would help them automate and speed up their Supplier Invoice and Contract processing.

Solution: Haulfryn Group implemented PSICapture along with SharePoint to create a fast, automated document capture solution. PSICapture imports scanned documents and uses Advanced Data Extraction to index invoice and contract information. The Invoices and Contracts are then automatically published to the appropriate SharePoint library for approval and payment.

Mission Linen Supply

Problem: Mission Linen Supply had more than 35,000 contracts at their 45 locations and lacked a centralized system where they could be stored and easily retrieved. Outsourcing was considered, but the company ultimately realized they needed an in-house solution so that new contracts could be easily scanned in at each location and published to the centralized repository.

Solution: Contracts are scanned in daily from every location. They are indexed at one of three PSICapture workstations, then published to multi-level folders in SharePoint Document Libraries. The system allows for improved visibility and forecasting through document automation, as well as enhanced reporting capabilities through PSICapture's intelligent indexing and ability to fully integrate with SharePoint.



Bringing it All Together

Using SharePoint should be a simple and efficient way for your business to manage information. It is designed to be a tool for storing and collaborating on files, reducing your business's reliance on paper documents. The operations you can complete with SharePoint are extremely robust—from organizing files and sharing ideas to building company intranets and external websites.

The important thing to remember is that SharePoint is only as useful as you set it up to be. It can only function to its full capacity if you enable it to do so by taking advantage of features like Managed Metadata and Content Types. And the best way to do that is to utilize PSICapture to accurately capture your documents, convert them into searchable PDFs and seamlessly route them to your SharePoint site.

PSIGEN offers a wide range of options for implementing a SharePoint solution. Whether you have a small business only requiring a small amount of scanned documents or you run a large corporation with extensive scanning needs at multiple sites, PSIGEN can help create a custom solution to meet your needs, at a price point that's unmatched for the value.

If you want to discover how to put PSICapture to work for your business, contact a PSIGEN sales representative.



PSI**capture**



SharePoint