

## Image API Recommendations

# Document Conversion and Digital Archival

Document conversion is one step in transitioning information assets from paper or film to digital files. The other equally important step is archiving digital data in a way that will preserve the information well into the future while enabling the organization to access the information virtually instantly on a daily basis.

Image API offers this road map for organizing, collecting, and presenting information in a project design plan for document conversion and digital archiving. The road map is based on our 16 years of experience providing document conversion and archival solutions to state and local government.

We understand that most groups do not procure these types of services and solutions on a regular basis. Therefore, a level of frustration with the “we don’t even know what questions to ask since we don’t know the issues” can occur. What we have attempted to provide is a fairly detailed road map that identifies important aspects to include in an RFP, pitfalls to avoid, and to make you aware of certain “standards” that have been developed that may help you to ensure that the work done or solution provided meets certain minimum standards of quality. These standards also may enable you to compare deliverables based on measurable outcomes. We hope this is helpful.



# Roadmap to Success

A successful document conversion and digital archival project requires careful planning and execution. Planning should be detailed and include a thorough examination of the organization's goals for the conversion and digital archival. Good planning, with the right measures to ensure accountability, will drive efficient execution and ultimate success.

Most organizations will have overall goals for transitioning to digital data. Common examples include improving work processes, enabling secure document storage, and archival cost savings; however, more specific information will be needed to fully execute on the plan. Organizations should take into account near term and long term use of digital documents. They should consider their needs for image and data quality, security, storage and retrieval, legal document preservation, timeliness, cost, and vendor credibility. They will need a firm understanding of the documents to be converted, groups that need to use them, technical infrastructure and support, desired vendor capability/qualifications, and more.

Every element of these requirements should be documented and explained as accurately as possible in order to drive accurate bid pricing and lay the groundwork for a disciplined approach to the project by the chosen vendor. Clear, detailed planning will set realistic expectations and improve the prospects for success.

Key information and/or requirements in the following areas should be gathered and communicated:

1. Document Population
2. Security Requirements
3. Chain of Custody/File Tracking
4. Quality Assurance
5. Vendor Standards
6. Project Management
7. Document Conversion Process
8. Experience and History
9. Vendor References

# Document Populations

It is extremely important to the vendors to be as clear as possible when you describe the collection to be converted. Vendors rely on a thorough understanding of the client's document population to provide an accurate price estimate and detailed project plan.

Factors such as volume of documents, variations of document types, document sizes and conditions, indexing requirements, QA levels, and similar characteristics affect all aspects of our services, particularly the price. And, this is also the area where vendors will cut every corner they can to reduce their price and then later request a change order to raise the price once they have won the business. This is one of the key pitfalls, not being detailed enough in describing your collection. (Note: Image API audited all our projects over 15 years and found less than 1% of had a change order.)

We recommend collecting the following key information about your document population:

- Total number of pages to be converted (remember a two page document equals two images so make sure you inform people of how many images you estimate you have or at a minimum estimate how many pages might be double sided)
- Number of folders (or files). Average number of documents within each folder
- Indexing
- Level of indexing required
- At the folder level i.e. First Name, Last Name, Acct number or
- At the subset level i.e. Folder level, then Car Loan, Mortgage, etc. or
- At the page level i.e. Folder level, Car loan, Loan application.
- The location where the conversion is to take place (onsite or offsite)
- Paper type and conditions of documents
- % one-sided; % two-sided
- % color; % black and white
- Scanning DPI (dots per inch; usually 200 or 300 dpi)
- Staples? Paper clips? Sticky Notes?
- Document stored: in folders? in binders?
- Availability of an existing index database (if database, what is the primary key)
- Average or exact number of characters in each index
- Location of indexes on documents
- Quality Assurance level required
- Output file format
- Delivery media (CD, DVD, SFTP, etc.)
- Do documents need to be refastened or can they be put back in loosely.
- Are documents to be destroyed after the digital images have been approved or do you want your files returned

Additionally, good vendors will use the information provided to identify opportunities to reduce the client's costs. For example, we can index document images using a manual method, an automated method, or OCR technology. Each has advantages and disadvantages mostly around cost, reliability, or availability of supporting databases. Having a good understanding of the client's documents, along with the organization's goals, mandates, budget, associated items i.e. an associated data base, and other constraints, gives us an opportunity to suggest a best practices approach.

Indexing is one of the two most costly aspects of converting documents so you need to determine how quickly you want your users to:

1. Find information
2. How large each file is i.e. how many pages per folder
3. And how frequently people access the information

The only time we suggest you index past the folder level is when the files are heavily used, have a good number of pages in each file (25 or more as a rule of thumb), and quick access will have a big impact of customer service levels. Otherwise, indexing at the folder level will enable you to find your document quickly at a fairly low cost.

# Security Requirements

Security for client documents and the information they contain should be as important if not more important than the quality of the deliverable. The security of the content on the documents such as SSN, addresses, etc. normally determines whether the work is done onsite (because the information is too sensitive to allow it to leave the facility) or offsite. Regardless, in all of our work with state government, security is a requirement, and chain-of-custody is paramount.

Generally, security includes safeguarding the physical documents, protecting the integrity of information in the documents, protecting the confidentiality of information, and providing disaster recovery planning for the digital data and images.

Key requirements for security should include the ability to adequately describe the following:

- Procedures to assure each file is logged out and accounted for throughout the process until it is logged back in or verified that it was destroyed
- Procedures to ensure the safety of physical documents in transit and in production
- Procedures for protecting the integrity of information contained in the documents
- Backup and disaster recovery options
- Hiring policies and security requirements for personnel
- Facility security
- Health Insurance Portability and Accountability Act (HIPAA) qualification to meet federal protections for personal health information held by covered entities

We recommend you request a written set of security measures and practices to protect your records during all phases of document conversion and digital archiving.

We further recommend the following:

- No Sub-Contracting or offshore Processing of Images and/or Data
- PMI Certified Project Managers
- Certified Document Imaging Architect
- Level One Background checks on Document Processing Staff
- Biometric access control for Zoned Employee Access
- Intrusion detection system
- Onsite or offsite services options

- Web-based Tracking System for Boxes and Records Moving Through Process
- Security Vault for Storage of Sensitive Records
- References of Handling Confidential Information
- Tour Facility Prior to Award (This is very important and should required)
- HIPAA certification and annual training of employees
- Image API warrants under the Security Rule of HIPAA that our document management software can incorporate the functionality that satisfies requirements for vendors supplying applications and solutions under the technical safeguards in the Technical Category.

# Chain of Custody/File Tracking

Closely related to security is the ability to track files and to ensure secure transport of documents with chain-of-custody handling procedures. Tracking also is important if the files are in active use. What happens when you need a file and it is offsite being converted? You need to be able to send a “file request” and have that file returned digitally to you within your specified time frame to meet your business requirements.

For example, we provide a File Tracking System as part of the overall suite of software applications Image API relies upon when performing file conversion/management projects. It is unique in that it is one of the few applications that is also used by our customers either as part of an Image API project, or as a standalone application to help them manage their files internally. The File Tracking System was designed to address the “chain of custody” aspect of each conversion project.

Chain of Custody is a legal term referring to the procedures followed when documents or evidence moves from one place to another. Although the procedures may vary from project to project, it is imperative that chain of custody procedures be documented and followed for each File conversion/management project.

As with most software applications, a file tracking system is a tool to help implement an overall policy or procedure. For example, the procedure may be a “Backfile Conversion Procedure” which would address items such as the following:

- Definition of the group of files to be converted as part of the project
- Definition of the order of file conversion, e.g. alphabetically, by case number, by year, etc.
- Definition of a unique identifier for each file, e.g. name, case number
- Procedures to check-out each file from the customer site
- Procedures to check-in each file to the Image API processing site
- Requirements to track files as they move through the different phases of conversion, e.g. document prep. and quality assurance, scanning and quality assurance, data entry and quality assurance, output and quality assurance
- Definition of how to handle requests for documents that are needed by the customer while in Image API’s possession
- Procedures to check-out each file from the Image API processing site
- Procedures to check-in each file to the customer site
- Procedures for when the customer receives documents to be added to a folder that is checked out to Image API
- Procedures for when the customer receives documents to be added to a folder that has already been converted by Image API
- Definition of the customer’s plan to perform quality assurance on the folders converted by Image API

It is a mark of the file tracking application's flexibility that it can be used not only in backfile conversion projects, but in any file management project. For this reason, we make the File Tracking Application available to customers as a standalone system used to manage the movement of their files internally within their organization.

# Quality Assurance

When all said and done, the ability to instantly retrieve and clearly read the information is job one when converting files.

Key indicators of quality should include:

- Documented QA processes
  - Document handling
  - Document prep
  - Document scanning
  - Document indexing
  - Document QA
  - Document output
- The legal standard requires random sampling of imaged documents. This is a process we following in all of our document processing services. It will help protect the State against challenges to the legality of imaged Social Services documents.
- Standards based policies and procedures
  - Adherence to industry standards
- Work references related to quality

You should require a Quality Adherence Manual to determine if the vendor runs its operation in a proper manner or is a five- to ten-person shop with old equipment that will quote a low price but not be equipped to do large volume work.

Sections from the Image API Quality Manual include the following:

- Document Production Flow for ANSI standards, image quality, file integrity, indexing accuracy, and completeness
- Document Prep production job standards, job duties, ANSI/AIIM standards
- Document Scanning production job standards, job duties, ANSI/AIIM standards
- Indexing Data Entry production job standards, job duties, ANSI/AIIM standards
- Quality Assurance production job standards, job duties, ANSI/AIIM standards
- ANSI/AIIM Technical Standards

# Vendor Standards

The value of following industry standards in document management cannot be overstated. **Key point to know and major pitfall: If the vendor does not scan to ANSI standards, the images may not be legal documents.** Standards are designed to ensure that digital documents remain as useful and as legally valid as paper. The standards help ensure high quality document images, retrieval, and storage well into the future.

The American National Standards Institute (ANSI) is the organization responsible for setting the standard by which the document imaging industry operates. ANSI standards provide a level of comfort in two areas: the quality of the images and the legality of the images. ANSI Standards during the document conversion process is written into code in many states and is part of the federal requirements for capturing, storing, indexing, and maintaining electronic records.

Image API follows ANSI standards to meet legal requirements in document processing. The legal standard requires random sampling of imaged documents.

Key ANSI standards relative to document processing include:

- ANSI/AIIM TR25-1995 – The Use of Optical Disks for Public Records
- ANSI/AIIM TR27-1996 – Electronic Imaging Request to Proposal (RFP) Guidelines
- ANSI/AIIM TR28-1991 - The Expungement of Information Recorded on Optical Write-Once-Read-Many (WORM) Systems
- ANSI/AIIM TR31-2004 – Legal Acceptance of Records Produced by Information Technology Systems
- ANSI/AIIM TR32-1994 – Paper Forms Design Optimization for Electronic Image Management (EIM)
- ANSI/AIIM TR 33-1998 – Selecting an Appropriate Image Compression method to Match User Requirements
- ANSI/AIIM TR34-1996 – Sampling Procedures for Inspection by Attributes of Images in Electronic Image Management (EIM) and Micrographics Systems
- ANSI/AIIM TR35-1995 – Human and Organizational Issues for Successful EIM System Implementation
- ANSI/AIIM TR40-1995 – Suggested Index Fields for Documents in Electronic Image (EIM) Environments
- ANSI/AIIM MS52-1991 – Recommended Practice for the Requirements and Characteristics of Original Documents Intended for Optical Scanning
- ANSI/AIIM MS53-1993 – Recommended Practice; File Format for Storage and Exchange of Image; Bi-Level Image File Format: Part 1
- ANSI/AIIM MS55-1994 – Recommended Practice for the Identification and Indexing of page Components (Zones) for Automated Processing in an EIM Environment

# Project Management

Project management plays an important role in large scale conversion projects and software implementations, both on the client side and the vendor side. Document management is a specialized field, and for this reason it requires project management talent with specific experience. Project managers should have a technical background as well as experience in solving business issues. The ability to apply technology to work processes is highly valuable.

Key components in assessing a company's project management capabilities should include the following:

- Project manager resumes
- Technical background
- Years of experience on document conversion and document management software technology projects
- References on the company and on the project manager
- Information on large scale conversion projects both on-site and offsite
- The project manager's interpersonal and communications skills
- Experience or examples of skills in issue resolution (every project has issues, large or small; successful resolution is vital to the overall success of the project)

Professional project management is a critical success factor in any project. We leverage the techniques taught by the Project Management Institute, our vast experience supporting clients, and knowledge of industry best practices to ensure mutual success.

# Vendor Document Conversion Process

Vendors should be required to describe their document conversion process. The following is a brief example of ours process. Companies should be able to describe their processes when submitting bids.

## **Document Conversion Process**

The document conversion process is designed to ensure consistent, high quality, and secure results. It follows the five general steps of document preparation, document scanning, image enhancement, indexing, and output.

## **Document Preparation**

Our document preparation process begins upon receipt of the project's documents at our production facility. Documents can arrive via any number of methods, including Image API transport, customer transport, or commercial transport (FedEx, UPS, DHL, USPS, and others). Image API employs chain-of-custody practices and HIPAA compliance when handling confidential documents.

The document preparation process follows the steps below:

1. Receive and inventory all containers.
2. Reconcile to shipping manifest.
3. Generate unique barcode and apply it to each container.
4. Insert section identification separator sheets to maintain record, folder, document, and document type integrity.
5. Prepare individual documents for scanning by removing paper clips, staples, binder clips, rubber bands, and other fasteners from pages.
6. Move "Post-it" notes to open areas on the page or onto a clean sheet of paper to avoid covering information.
7. Repair torn or damaged pages and straighten all folds.
8. Tape documents smaller than 5 1/2" by 8 1/2" to standard-sized sheets of paper. Tape leading edges.
9. Scan container barcode label to enter the container into Image API's DTS tracking system.
10. Move container into appropriate pre-scanning storage area.

## **Document Scanning**

Image API uses state-of-the-art production scanners capable of high volume and high quality throughput. Our document conversion specialists follow a scanning process designed to operate at peak efficiency and produce precise images.

The document scanning process follows the steps below:

1. Register scanning job into DTS tracking system.
2. Scan container barcode ID to identify start of scanning.

3. Enter employee ID for quality and chain-of-custody tracking.
4. Ensure scanner operation by checking scanner setup, optics, and calibration.
5. Ensure scanning job readiness.
6. Check for appropriate document orientation.
7. Insert documents into automated feeder.
8. Begin scanning.
9. View all scanned images on the scanner workstation to ensure quality.
10. Retrieve scanned documents and place them back into the container.
11. Register end of scanning job in DTS tracking system.
12. Scan container barcode ID.
13. Enter employee ID.

### **Image Enhancement**

After documents have been scanned and checked for completeness, Image API employs an image enhancement application that automatically enhances image quality and increases the recognition accuracy of any automated process. The application performs the following functions:

- Straighten skewed and angled images.
- Detection and removal of black borders.
- Removal of lines on typewritten forms.
- Removal of random pixels in image (Despeckling).
- Removal of shaded backgrounds.
- Sharpen the edges of type characters (Edge enhancement).

Image API offers image recognition services to enhance the usefulness of scanned documents. The five primary types of image recognition are as follows:

- Form ID recognizes different types of forms. For example, it can separate a patient's medical records based on form type: admission assessments, pathology reports, anesthesia records, etc. Form ID is especially useful in situations where pre-printed forms are widely used.
- OCR (Optical Character Recognition) recognizes machine printed characters. It is popular for keyword searches. OCR is very good, but rarely achieves 100% accuracy.
- ICR (Intelligent Character Recognition) recognizes hand printed characters. ICR is less accurate than OCR due to wide variability in individual writing styles.
- OMR (Optical Mark Reading) recognizes check boxes, filled-in bubbles, etc. and is highly accurate on properly marked forms. Typically, OMR is used to automatically tabulate the results into a database (such as for surveys or voting).
- Barcode recognition extracts information from pre-printed barcodes. It is designed to automate the capture and processing of a wide variety of data. Barcode recognition is highly accurate.

Image API is skilled in the full range of image recognition software and can recommend a program to best meet the accuracy requirements of the most demanding applications.

### **Indexing Expertise**

Image API applies accurate and consistent indexing to all digital documents to ensure reliable access to them after scanning. Accurate indexing is absolutely critical to the success of any document conversion project: It allows you to quickly retrieve documents using several different search criteria. We can index by any number of elements and can automate the process using existing databases supplied by the customer.

Image API offers multiple recognition processes (OCR, barcode, etc.) to index documents automatically, or in cases when that is not reliable or practical, can manually enter indexing data. We validate and verify all indexing data, offering all methods:

- Independent double key entry - Two different trained operators enter data at different times while looking at the same scanned image. If the entered data does not match up perfectly (spelling, spacing, etc), software rejects the data entry and it is routed (along with the image in question) to a supervisor who resolves the difference. In this manner, every index field is verified at least twice for its veracity. This process helps ensure 100% accuracy.
- Database validation – A customer database is queried upon the entry of an index field, which retrieves other elements that allow the operator to verify the accuracy of the index. For example, instead of manually entering name, SSN, date of birth, and address for every document, the customer supplies a database that allows us to enter a unique identifier (SSN) and pre-populate the remaining fields. Database validation increases accuracy and decreases cost.
- Audit controls – Auditing helps ensure that all scanned documents get indexed only once and tracks the personnel performing data entry.
- Persistent Indexing – This method enables the system to automatically remember the last set of index attributes used in cases where all documents in a batch belong to the same person, or where all documents in a batch belong to the same document type.
- Barcode separators – This method enables the design of barcode separator sheets that contain all of the indexing data required to identify a document and its contents (document typing).

Image API separates the indexing and validation functions from the scanning function. This allows multiple personnel to simultaneously index the output of a single high-speed scanner, which results in cost and time efficiencies.

## **Output**

Image API can output files to any format desired and deliver them on any media as specified. The most common file formats are TIFF, PDF, and JPEG, but Image API can output to all proprietary and non-proprietary industry standard formats as well.

Image API will deliver image files via CD, DVD, web, FTP, VPN, or external hard drive. Many customers will load their image files onto a central server and use CDs or DVDs as backup for archival storage and disaster recovery. Image API will produce the CDs or DVDs (with customized labeling and self-contained retrieval applications) and keep a duplicate copy in an archival vault for safekeeping, if desired.

Since many customers already have document management software, Image API is often called upon to import the completed image files into the customer's document management system. Our programmers and database administration personnel write load scripts, determine proper file structures, and generally ensure that everything works as the customer desires. Image API programmers are highly skilled in Oracle, SQL, and Access databases, as well as Visual Basic, .Net, XML, ASP, Java, and HTML programming languages.

# Vendor Experience and History

Experience and history are the best indicators of a company's abilities to provide professional, standards-based excellence in converting government documents and enabling reliable digital archival.

Image API is a government support company providing document management technology solutions, document conversion services, and business process outsourcing.

Our software and technology solutions help government automate and improve their information-intensive processes. Image API's flagship content management system, iCenter™, provides the foundation of many of our solutions to content centric business processes. Image API and iCenter have helped government agencies transition to digital content and improve productivity, improve services to constituents, reduce costs, and ensure their abilities to meeting regulatory/legal requirements. Our mission is to exceed our customers' expectations in every respect and to deliver true, benefit-driven solutions.

Image API operates two primary facilities; a headquarters center in Tallahassee, Florida, and a second center in Huntingdon Valley, Pennsylvania near Philadelphia. Both centers provide a full range of technical, project management, and document conversion services.

In addition to implementing large scale document management technology solutions for government clients, Image API converts millions of documents every year for state government from our production centers and on-site. We also offer complete document processing center service operations for states wanting to centralize document imaging to gain certain economies of scale. We work with all agencies and all document types, including irreplaceable historical documents, oversized engineering documents, and high confidential health, financial, and personal documents.

## Extensive Experience in Government

Image API has 16 years of knowledge and experience in providing document management solutions and converting documents to digital images specifically for government clients. In order to be successful for so long, Image API has developed a number of distinguishing characteristics:

- Technology, processes, professional project management, and highly skilled technical staff.

---

### Image API at a Glance

- 16+ years experience
  - Headquarters in north Florida, offices in Pennsylvania, Texas, Connecticut, South Florida
  - Strong financial rating by Dun & Bradstreet
  - Single-source solutions in document processing, software, and process improvement
  - Proven track record
  - [www.imageapi.com](http://www.imageapi.com)
-

- An extensive collection of state government customers and positive references.
- An unparalleled record of success—we have never had a project failure.
- Two highly secured production facilities protected by intrusion alarm detection, biometric access control (to prevent unauthorized access to key areas within the buildings), fire alarm systems, and disaster recovery procedures to protect the digital assets of our customers.
- Employee background checks and annual HIPAA training.
- An on-staff Certified Document Imaging Architect (CDIA).
- Adherence to ANSI and AIIM industry standards for document conversion and quality assurance.

### **Projects Similar in Volume and Time**

Image API holds long-term contracts with a number of state agencies in which we continually scan, index, and load millions of documents routinely throughout the year. We also have performed large scale “one-time” imaging projects for customers. All were performed on time and within budget.

# Vendor References and Project Descriptions

Most companies can provide a few references even if they have a history of poor performance. Some of the new communication tools i.e. LinkedIn and industry chat rooms can help you learn if another state encountered problems using a vendor. One of our most impressive boasting rights involves our 16-year record of 100% successful projects. We have never experienced a project failure.

In addition to references, vendors should provide brief project descriptions to quality themselves as eligible for consideration.

Key information from references should include:

- State government references
- Client name, address, phone number
- Contact name, phone number, email
- Project description and relevance to the client's project
- Size of the project

Project Descriptions should be brief, but adequate to illustrate eligibility.

# A Partial List of Image API Services & Solution Offerings

## **Document Conversion Services**

- Mailroom
- Document transport
- On-site, off-site
- Paper (all types) to digital
- Standard, oversized documents
- Microfilm/microfiche to digital
- Document Processing Facilities Management
- Project Management
- Process reengineering
- Document preparation
- Scanning – b&w, color, 200 – 1200 dpi, single side/duplex
- Data entry, automated, OCR
- Quality Assurance according to ANSI/AIIM standards
- File output to any format
- Data delivery via SFTP, CD, DVD, hard drive

## **BPM Solutions**

- Business Process Consulting
- BPO Services
- iCenter ECM System
- Vital Records Automation
- Infectious Disease Case Management System
- Travel Request/Reimbursement System
- Board Agenda and Hearings