Glossary of Scanning Terms

Archival Storage

Archival storage involves the safe and effective storage of documents for long periods of time in a manner in which the documents may be retrieved when needed, but not necessarily quickly (possibly even overnight). Key concerns include cost, amount of storage space required (physical or digital), long-term viability of the storage medium, security, accessibility and the long-term financial and operational stability of the supplier.

Off-line Storage includes storage of data on computer tapes, CDs, DVDs and Magneto-Optical drives that are not mounted and accessible at all times. This includes the use of off-site facilities such as hardened silos or warehouses to store such devices. This is typically used for documents that must be maintained for regulatory reasons, but are not expected to be accessed again, or would be accessed only under highly unusual circumstances.

Near-line Storage includes the use of automated tape silos or libraries that can mount the storage device with a short delay. Near-line storage is used for documents that are rarely accessed, but may still be considered active.

Backfile Conversion

See **Production Scanning**

Business Continuity

Business continuity plans are used to ensure that a business or other organization can continue to function in the wake of a disastrous event such as a fire, flood, earthquake, etc… While natural disasters and fires are frequently considered, thorough planning should include events such as the loss of key personnel (through illness, death or separation from the organization) and key suppliers or partners as well. Off-site backups, hot sites (locations with adequate computer and office space for operation to resume), communications chains and other elements are parts of a complete business continuity plan.

Disaster Recovery is the principal component of a business continuity plan.

Business Scanning

Also known as business imaging, business scanning involves the use of scanning as part of a repeatable process used by a business or organization in its day-to-day operation. Examples include:

- Scanning invoices when they arrive at the mail room so that they can be processed and paid.
- Scanning receipts to attach electronically to an expense report
- Scanning contracts for long-term storage in a secure network location
- Scanning checks for processing by a financial institution
- Scanning photographs to store with an insurance claim record in a database

Disaster Recovery

See **Business Continuity**

Document

- A document is any collection of information an organization uses to conduct its affairs. Documents may be in almost any format, including paper, film, digital computer files or digital images, even audio files or recordings.

Document Life Cycle

A Document Life Cycle consists of all the various stages a document goes through from creation through final destruction. The specific phases of a given document life cycle will depend on the document and the business process associated with the document. They can include creation, editing, distribution, archival, retrieval and destruction, among others. An example of a document life cycle for a paycheck might be:

1. Paycheck generated by the payroll system
2. Reviewed by staff accountant
3. Rejected for editing by payroll manager
4. Revised by staff accountant
5. Approved by payroll manager
6. Signed by treasurer

7. Copy of paycheck printed and mailed to employee

8. Electronic copy of paycheck stored in near-line archival storage for 90 days

9. Electronic copy retrieved and printed for audit

10. Electronic copy transferred to off-line storage for 7 years

11. Electronic copy deleted.

Document Management Systems

Document Management Systems are large databases in which documents are stored, along with fields of information which enable quick and easy search and retrieval of those documents. Document Management Systems often have features to support automatic implementation of document retention policies and security measures to ensure only authorized individuals can access each document. They may also provide automated distribution or step documents through a workflow that routes documents to the appropriate party at each step. Another common feature of Document Management Systems is document versioning, which maintains a historical record of the document, while ensuring that all users see the most current version by default.

Document Retention

See **Regulatory Compliance**

Intelligent Character Recognition (ICR)

See **Optical Character Recognition (OCR)**

Metadata

Metadata (literally data about data) are pieces of information used by document management systems to make searching for documents in the system easy and effective. Some examples of metadata fields include:

- Date of Creation

- Document Author

- Account Number

- Document Type (for instance: Invoice, Receipt, Bill of Materials…)

- Document Status (for instance: current, expired, draft, released…)

Magnetic Ink Character Recognition (MICR)

See **Optical Character Recognition (OCR)**

Near-line Storage

See **Archival Storage**

Off-line Storage

See **Archival Storage**

Optical Character Recognition (OCR)

Optical Character Recognition, or OCR, is a process by which a digital image of a paper document is translated into digital text that can be edited by programs such as a word processor. OCR is an inexact science, affected by factors such as the condition of the original document, the quality of the digital image, the orientation of the text, the font size and type. All OCR processes experience errors. For instance, the letters **o**, **e** and **a** all appear somewhat similar and are approximately the same size, leading OCR systems to make errors if the quality of the original image is poor. Some types of writing, especially handwriting, can be exceptionally difficult for computers to interpret. Several measures have been developed to address these issues.

MICR is a type of font and magnetic ink (or toner) that helps to minimize OCR errors. The design of the characters along with the magnetic properties of the ink used ensure a very low error rate for MICR readers. MICR ink and characters are used specifically for checks and other financial documents that are handled by banks.

Zonal OCR is a method of converting only specific regions of a document to text. Zonal OCR is used on forms, where the bulk of each document is identical and only information in certain fields needs to be stored. This can significantly reduce the storage requirements for the resulting document, and also limit the number of errors, since much of the document is not read.

Intelligent Character Recognition (ICR) is an advanced form of OCR which attempts to convert handwriting to digital text. ICR frequently makes use of the expected context to interpret the text (for instance, a zip-code field should only contain the digits 0 through 9, so the target character should not be interpreted as the letter Q). Postal address interpretation is one of the principal uses of ICR currently.

Production Scanning

Production scanning is a type of business scanning involving large quantities of documents processed by dedicated, sometimes high-volume (more than 100 images per minute) scanners, often in batches. It frequently involves the storage of the resulting document images in a database (such as a Document Management System) with additional information, known as metadata.

A **Backfile conversion** is a type of production scanning where large quantities of paper documents stored for many years are scanned and stored electronically in a one-time conversion in order to reduce storage costs and improve accessibility and security of storage. Key considerations for backfile conversion include the cost of leasing or purchasing scanning equipment, the number of people and time required for the conversion, the quantity and type of digital storage media required, the data fields necessary for search and retrieval of the documents, and the physical format and condition of the original documents.

RA2K

RA2K is the hardware architecture of most Ricoh MFP equipment that allows applications to be installed on the MFP's hard disk. RA2K supports two types of applications, C and Java. Applications include web browsing applications, new printer definition languages and communications applications.

Regulatory Compliance

Regulatory compliance involves the maintenance and retention of documents according to specific rules set up by an outside agency. Businesses are subject to regulatory compliance based on the laws of the country or state in which they conduct business, such as the Sarbanes Oxley act, which regulates the maintenance of business records or the Health Insurance Portability and Accountability Act (HIPAA) which regulates access to medical records by authorized personnel. They may also be subject to voluntary regulatory compliance in order to achieve and maintain certifications based on international standards such as ISO 14000, an international standard for environmental management or ISO 9000, which dictates documentation standards to ensure quality control.

Document retention requirements are typically driven by federal and state legal requirements that business, tax, employee, medical and student records be maintained for specific periods of time. Some regulations may specify the minimum length of time records are retained; other may specify the maximum period that records may be retained.

Solution

A solution is a product or service provided by one company to assist another company in processing or maintaining documents through one or more stages of their document life cycle. Solutions may be computer software, hardware, on-line or manual services or even business processes designed by one company for use by another. Some examples of solutions are:

- A software system that stores documents in a database with attached data fields to make search and retrieval easier

- A hardware system that tracks all print devices on a network and routes print jobs to the most cost effective available printer

- A service that provides staff to manage incoming mail, scan it to digital files and route it to the appropriate recipient(s)

- A business process that utilizes MFPs and various software products and online services to manage client prospecting, order generation and fulfillment and billing to minimize delays and ensure Sarbanes-Oxley (SOX) compliance of business data

Zonal OCR

See **Optical Character Recognition (OCR)**