

Session Title: Digital Transformation Starts at Capture: Converting Paper into AI-Ready Data

Session Description: Forward-thinking organizations recognize the untapped potential within their paper archives - both as valuable training data for AI and as a way to reduce physical storage costs. Yet the key to unlocking this value lies in getting data right from the start. This session reveals how intelligent capture workflows not only prevent costly downstream errors but transform decades of business information into AI-ready digital assets. By streamlining your capture workflow, you can validate data accuracy at the point of entry, eliminating expensive cleanup efforts while boosting productivity across your organization. Through practical examples, we'll explore proven capture solutions that address common business challenges like invoice processing, customer onboarding, and compliance documentation. By the end of this session, you'll have actionable strategies to harness the hidden value in your paper archives and build a capture foundation that delivers the clean, reliable data your AI initiatives need to succeed.

Digital Transformation Starts at Capture

Converting Paper into AI- Ready Data

April 2024



Poll Question: What industry segment are you from?

Government – (FED, ST, or Local)

Healthcare

Financial

Education

Manufacturing

Other



*"Study the past if you
would define the future"*

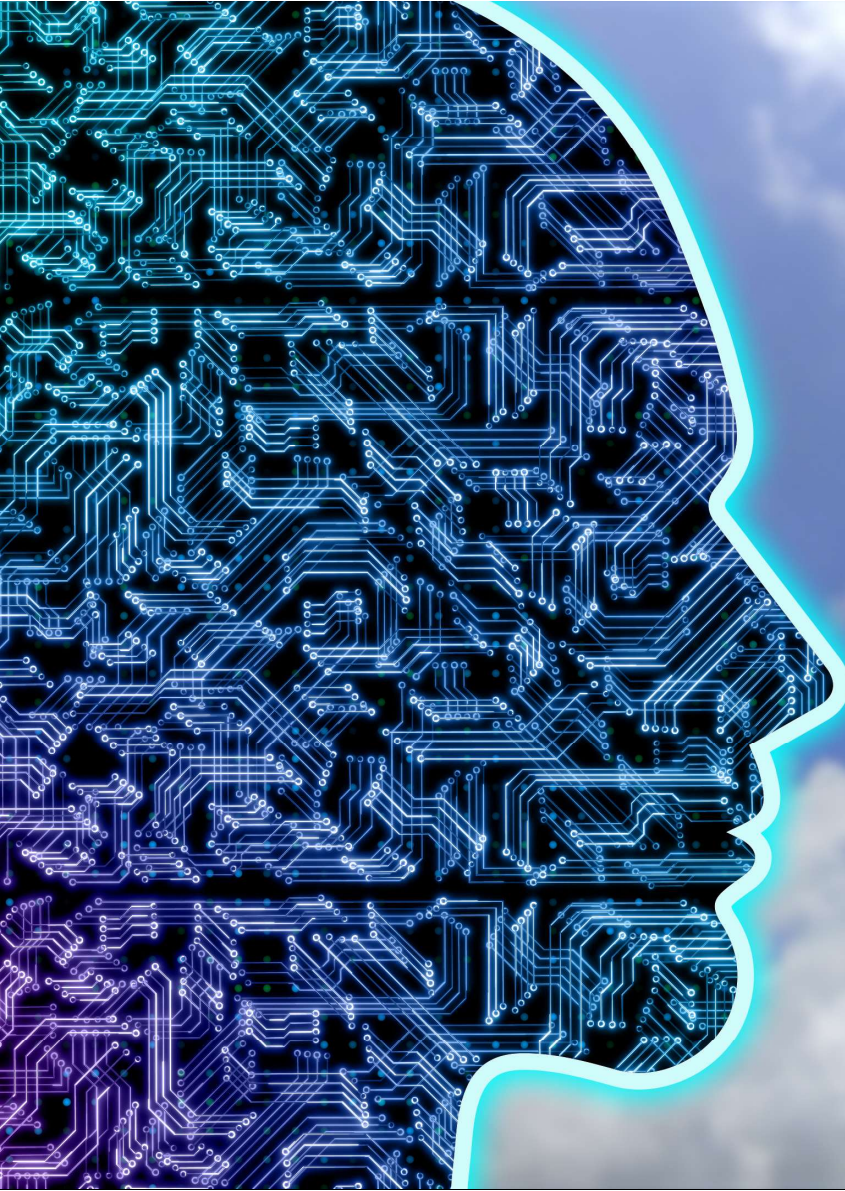
Confucius

c.5th Century BCE



The Promise of the Future...

Requires Data
(Accurate Data!)



Best Practices for Data Hygiene

Capture your data with the correct equipment & settings

Identify and remove redundant data from file archives

Classify & index your data

Implement metadata for access rights

Anonymize / Redact Sensitive Information

Chain of custody – Who, Where, & When, and How

Document Relevance/Ranking



Poll Question: Is your company or organization storing paper records?

Yes, we're storing them on-site

Yes, we're storing them off-site

Yes, we're storing them on-site and off-site

No, we're 100% Digital



The Cost of On-Site Document Storage



| City | Average Cost (sq ft) | File Cabinet 44" x 21" |
|---------------|----------------------|------------------------|
| San Francisco | \$61 | \$394 |
| Seattle | \$38 | \$246 |
| Chicago | \$34 | \$219 |
| Dallas | \$27 | \$172 |
| Atlanta | \$27 | \$173 |
| New York City | \$80 | \$514 |

What is the actual cost?

Storage

Retrieval

Risk

Insite

Source: Google AI Overviews / <https://www.loopnet.com/>

Poll Question: How much time do employees spend searching for data? (40 hour work week example)

4 hours

10 hours

20 hours

More than 20 hours



Poll Question: On average, what percentage of business decisions are based on data?

75-100% (Almost all of the time)

50-74% (Usually)

25-49% (Occasionally)

0-25% (Seldomly)



Is Your Company Making Data-Driven Decisions?

58% of respondents to a BARC survey¹ showed that managers are making decisions “On the cuff” or based on a “Gut Feel”

When your employees can't find data, they're left to their own devices to compensate.

This can cost your company money, drain productivity, and cause a loss of customer trust.

Not being able to find accurate data quickly is the top reason employees bypass it.

¹<https://barc.com/business-decisions-gut-feel/>





The Mystery of Dark Data

Data that you're storing, but not using

Unstructured and unclassified...and a potential value or risk to your organization

Includes data from employees who have left the company or moved departments

It could be 50% of an organization's data...or more. How much do you have?

How much of that data is on paper? In network file shares? In legacy systems?

Data is the key to future business success. Organizations that can harness dark data will have an advantage over their competition.

Poll Question: On average, how much does a data error cost an organization?

\$25 (About a dozen eggs)

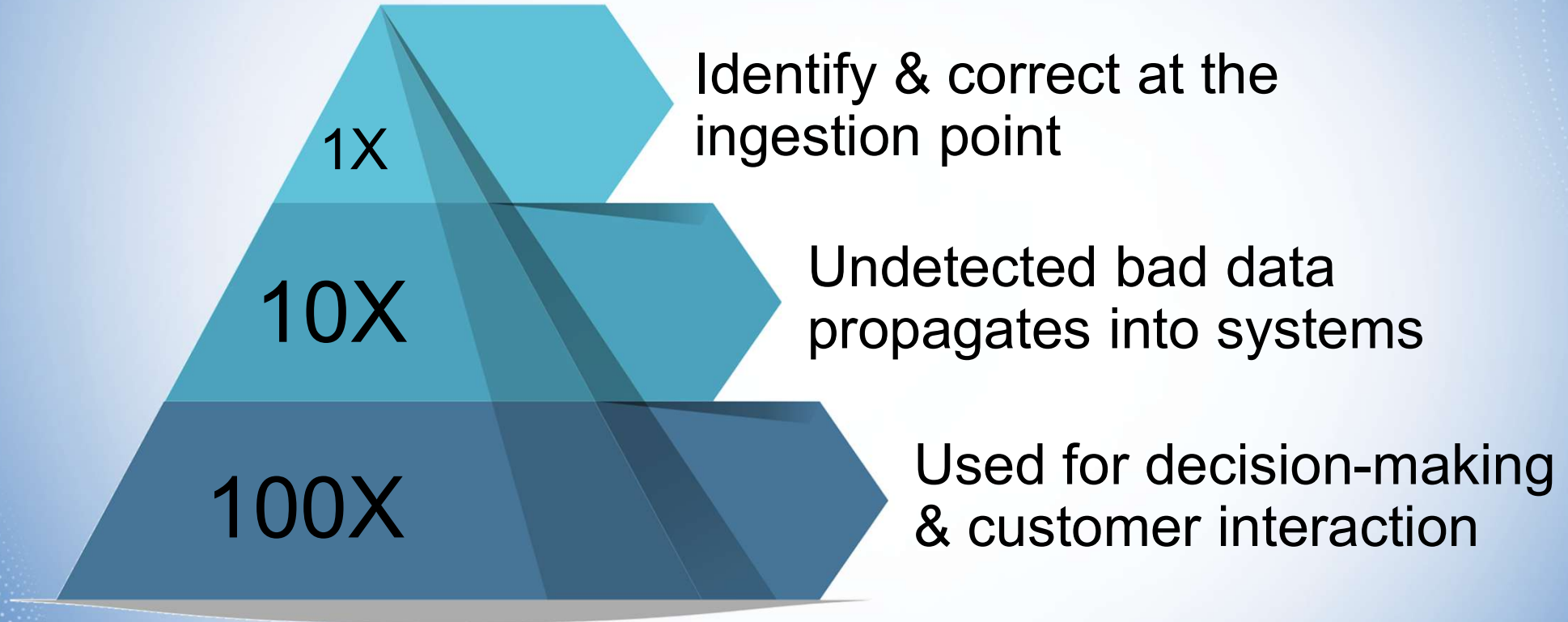
\$50 (Half tank of gas)

\$75 (2 drinks & a hot dog at the game)

100\$ or more!



The Rule of 1:10:100 - The Cost of Data Errors



The True Cost of Data Errors For Scanning Capture Workflows

Significant productivity loss, especially if the error is detected downstream

AI data hallucinations

Customer dissatisfaction and loss of trust / reputation

Regulatory fines for PII breach

Bad business decisions based on inaccurate data



“Conversion projects are labor intensive and take much longer than expected”



These are some of the challenges with document scanning workflows and archive projects



Pre- and post-scan processing is labor-intensive and costly



Manual classification and data entry waste time, resources, and are error-prone.



Forms are data-intensive and may have sensitive data



Workflow bottlenecks can stall information processing

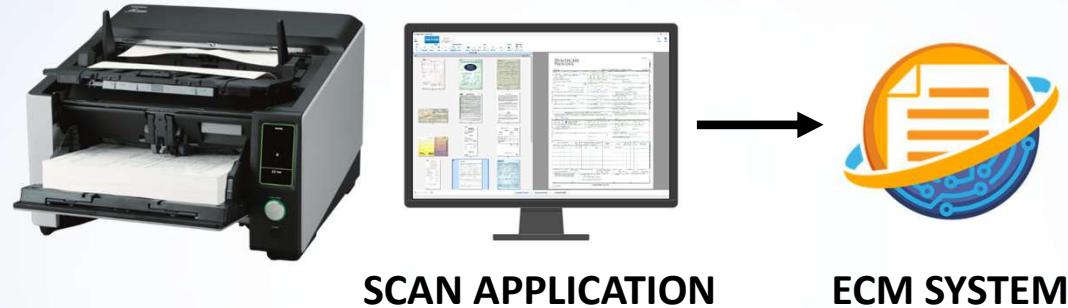


Capture solutions can be cost-prohibitive



Not all capture devices are optimized for high-quality image output and volume scanning

Capture Workflow Bottlenecks Can Stall Information Processing



- Too many steps for scanner operator
- Potential for errors
- Reduces throughput
- Hinders customer engagement



- 1 SCAN
- 2 QC
- 3 Index
- 4 Name File
- 5 Choose File Location

Poll Question: What type of devices do you use to scan paper documents?

Document Scanners

MFPs

Both document scanners & MFPs

None of the above



Components of an Intelligent Capture Workflow



Best in Class Scanning

Ricoh Document Scanners combine excellent paper handling and stellar image quality to ensure that documents are quickly and reliably digitized



Intelligent Image Processing

Automated document processing removes image artifacts and background noise to improve recognition accuracy & enhance the presentation of images



Advanced File Separation

Versatile multi-page document separation capabilities maximize scanner ADF capacity & overall throughput by reducing manual processing requirements.



Automated Data Extraction

Utilize a host of different methods to classify & index documents to extract information and accelerate retrieval in content management systems.



Effective Quality Assurance

Streamline QA workflow with an easy-to-use visual UI assisted by innovative features to highlight & correct difficult images.



Multi-Station Architecture

Scale the size, speed, and capacity of the capture solution by distributing the scanning, indexing, and QA workflow over multiple stations.

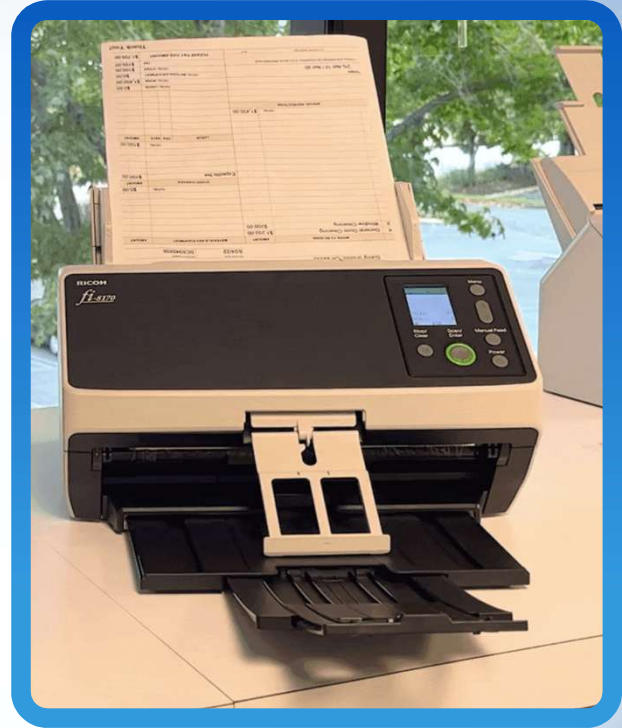


Seamless Integration

Multiple integrations with content management and cloud service solutions allow for fast and easy transfer of documents & metadata.



The Right Tool for the Right Job!



Paper
Handling



Image
Quality



Data
Integrity



Reliability

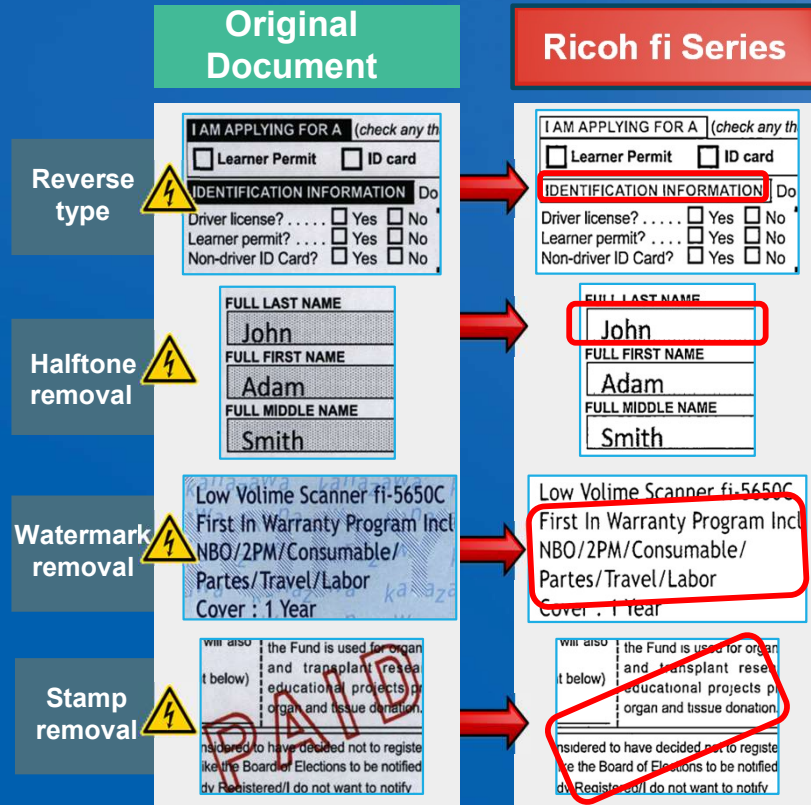


Easy
to Use



Automatic

Image Enhancement Produces Clean Data



AUTOMOBILE CLAIM REPORT
Agent uses

1

Billing Copy

184756382024 300

2

HEALTHCARE PROVIDER

3

City Courier

4

Billing Copy

184756382024 300

5

BILL OF LADING

6

XYZ Insurance

7

John Doe

8

Warranty Registration Card

9

BUSINESS REPLY MAIL

10

AUTOMOBILE CLAIM REPORT
Agent uses

11

Packing List

12

John Doe

13

DMV

14

REPORT OF TRAFFIC ACCIDENT

15

[illegible][illegible]

The image shows a sample of a "HealthCare Provider" form. The form is divided into several sections with various fields for data entry. The top section includes fields for "Patient Name", "Date of Birth", "Sex", "Race", "Ethnicity", "Religion", "Marital Status", "Social Security Number", "Home Address", "Home Phone", "Work Address", "Work Phone", "Email Address", and "Fax Number". The middle section is for "Insurance Information" and includes fields for "Insurance Company", "Policy Number", "Group Number", "Effective Date", "Expiration Date", "Premium Amount", "Deductible Amount", "Co-insurance Amount", "Out-of-Pocket Maximum", "Copay Amount", "Waiting Period", "Network Status", "Primary Care Physician", "Referral Required", "Pre-authorization Required", "Appeal Process", "Appeal Deadline", "Appeal Contact", "Appeal Phone", "Appeal Fax", "Appeal Email", "Appeal Address", "Appeal City", "Appeal State", "Appeal Zip", "Appeal Country", "Appeal Language", "Appeal Currency", "Appeal Units", "Appeal Frequency", "Appeal Duration", "Appeal Start Date", "Appeal End Date", "Appeal Status", "Appeal Reason", "Appeal Action", "Appeal Outcome", "Appeal Date", "Appeal Time", "Appeal Location", "Appeal Contact", "Appeal Phone", "Appeal Fax", "Appeal Email", "Appeal Address", "Appeal City", "Appeal State", "Appeal Zip", "Appeal Country", "Appeal Language", "Appeal Currency", "Appeal Units", "Appeal Frequency", "Appeal Duration", "Appeal Start Date", "Appeal End Date", "Appeal Status", "Appeal Reason", "Appeal Action", "Appeal Outcome", "Appeal Date", "Appeal Time", "Appeal Location". The bottom section is for "Medical History" and includes fields for "Current Medical History", "Past Medical History", "Family Medical History", "Social History", "Review of Systems", "Physical Examination", "Laboratory Tests", "Immunizations", "Vaccinations", "Screening Tests", "Diagnostic Tests", "Therapies", "Surgical History", "Anesthesia History", "Allergies", "Medications", "Vital Signs", "Vital Signs Date", "Vital Signs Time", "Vital Signs Location", "Vital Signs Contact", "Vital Signs Phone", "Vital Signs Fax", "Vital Signs Email", "Vital Signs Address", "Vital Signs City", "Vital Signs State", "Vital Signs Zip", "Vital Signs Country", "Vital Signs Language", "Vital Signs Currency", "Vital Signs Units", "Vital Signs Frequency", "Vital Signs Duration", "Vital Signs Start Date", "Vital Signs End Date", "Vital Signs Status", "Vital Signs Reason", "Vital Signs Action", "Vital Signs Outcome", "Vital Signs Date", "Vital Signs Time", "Vital Signs Location".

[illegible]

| | | | | | |
|---|--|---|--|--|--|
| 18475382024 | | 300 | | Billing Cycle | |
| From: <u>John Doe</u> <u>1234 Main St</u> <u>Anytown, NY 12345</u> To: <u>John Doe</u> <u>123 Main St</u> <u>Anytown, NY 12345</u> | | Payment Method <input type="checkbox"/> Credit Card <input checked="" type="checkbox"/> Debit Card <input type="checkbox"/> Bill Payment | | Billing Frequency <input checked="" type="checkbox"/> Monthly <input type="checkbox"/> Quarterly <input type="checkbox"/> Annually | |
| Service Name: <u>Basic Service</u> Service ID: <u>12345</u> Service Description: <u>Basic Service</u> Service Start Date: <u>01/01/2024</u> Service End Date: <u>12/31/2024</u> | | Payment Terms <input checked="" type="checkbox"/> Net 30 <input type="checkbox"/> Net 60 <input type="checkbox"/> Net 90 <input type="checkbox"/> Prepaid | | Billing Calendar <input checked="" type="checkbox"/> Standard <input type="checkbox"/> 30 Day Calendar <input type="checkbox"/> 45 Day Calendar <input type="checkbox"/> 60 Day Calendar | |
| Payment Method: <u>Debit Card</u> Payment Frequency: <u>Monthly</u> Payment Amount: <u>\$24.99</u> Payment Due Date: <u>02/28/2024</u> | | Payment Status <input checked="" type="checkbox"/> Paid <input type="checkbox"/> Pending <input type="checkbox"/> Overdue | | Payment History <input checked="" type="checkbox"/> 12 Months <input type="checkbox"/> 24 Months <input type="checkbox"/> 36 Months | |
| Payment Reference: <u>12345</u> Payment Description: <u>Basic Service</u> Payment Amount: <u>\$24.99</u> Payment Due Date: <u>02/28/2024</u> | | Payment Method: <u>Debit Card</u> Payment Frequency: <u>Monthly</u> Payment Amount: <u>\$24.99</u> Payment Due Date: <u>02/28/2024</u> | | Payment Status: <u>Paid</u> Payment History: <u>12 Months</u> Payment Calendar: <u>Standard</u> | |

| | | | |
|--|--|----------------------------|--|
| Name Students Unlimited, Pkwy, #11 | | Processing Date 6/17/73 | |
| Address DuPonts Unlimited, Pkwy, #11 | | City LITTLE ROCK, AR | |
| State Ark | | Zip 72202 | |
| BILL OF LADING PFU | | | |
| Date of Bill 6/17/73 | | | |
| Bill of Lading No. 545-555-1232 | | | |
| Consignee Students Unlimited, Pkwy, #11 | | | |
| Shipper Students Unlimited, Pkwy, #11 | | | |
| Commodity Students Unlimited, Pkwy, #11 | | | |
| Weight Students Unlimited, Pkwy, #11 | | | |
| Volume Students Unlimited, Pkwy, #11 | | | |
| Remarks Students Unlimited, Pkwy, #11 | | | |
| Signature Students Unlimited, Pkwy, #11 | | | |
| Date 6/17/73 | | | |
| Office Students Unlimited, Pkwy, #11 | | | |
| City Little Rock, Ark | | | |
| State Ark | | | |
| Zip 72202 | | | |

XYZ Insurance
Automotive Claims Div
123 Broadway Way
Chicago, IL 0346

12345
Date April 21 2013

NAME: Jon Doe
ID: 1234 56789

Three Thousand Twenty-Four and 87/100 DOLLARS

WENO #1234567 89

00006785 12345678

WENO 761-9

SAMPLE CHECK VOID NO VALUE

PFU

STANDARD FORM NO. 647

367

Jon Doe

DO NOT WRITE ABOVE OR BELOW THESE LINES
RESERVED FOR FINANCIAL INSTITUTION USE

FEDERAL RESERVE BOARD OF GOVERNMENT RESERVE BANK

| | | |
|---|--|--|
| Warranty Registration Card | | |
| Name <u>John Muga</u> | Date of Purchase <u>4/6/23</u> | <input type="checkbox"/> Muggies Update Line |
| Title <u>Office Manager</u> | Product Category <input type="checkbox"/> Hard Disk <input type="checkbox"/> Mouse | |
| Company Name <u>Acme Art</u> | Product Model Number <u>522023</u> | |
| Address <u>12545 First Street</u> | Product Part Number <u>52202</u> | |
| <u>San Francisco CA 94167</u> | Permitted Future product transfer (yes/no) | |
| Phone <u>415-555-2223</u> <u>415-555-2222</u> | Current region (Country) <u>USA</u> | |
| | Current region (Province) <u>Other</u> | |
| E-mail _____ | Location (City) <u>San Francisco</u> | |
| Department _____ | Operating System <u>Win XP Pro</u> | |
| Type of hardware <u>Art Supplies</u> | Interface <input type="checkbox"/> SCSI <input type="checkbox"/> IDE <input type="checkbox"/> Other <u>USB</u> | |
|  | | |
| Model <u>522023</u> | | Page 14 |

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 1000 BURNLEYVILLE, CA

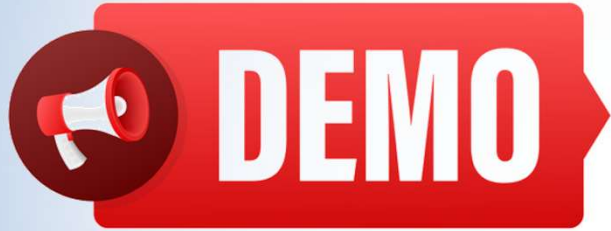
POSTAGE WILL BE PAID BY ADDRESSEE

MAL STOP 3-0
PFI AMERICA INC
808 E ARQUES AVENUE
BURNLEYVILLE CA 94005-0000

NO 004

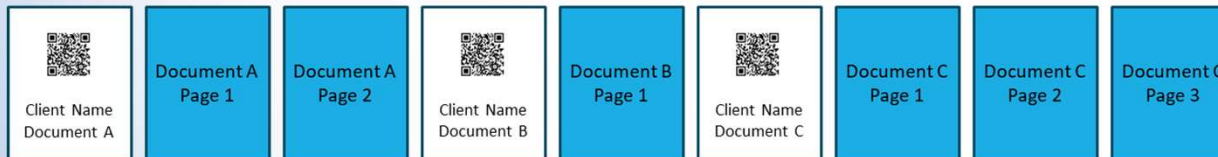
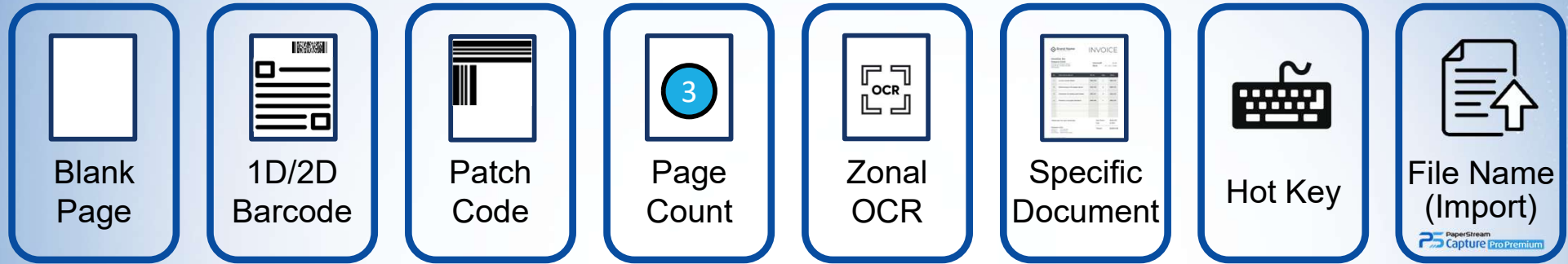
| AUTOMOBILE CLAIM REPORT | | | Page 1 of 2 |
|--|--|--|-------------|
| <div style="text-align: right;">04/07/2004 10:00 AM</div> | | | |
| Name: <u>John Smith, Jr.</u> Address: <u>123 Main St., Apt. 456</u> City: <u>Anytown, CA 90210</u> State: <u>CA</u> Zip: <u>90210</u> Phone: <u>(555) 123-4567</u> E-mail: <u>john.smith@anytown.com</u> | Name: <u>J. L. Smith</u> Address: <u>123 Main St., Apt. 456</u> City: <u>Anytown, CA 90210</u> State: <u>CA</u> Zip: <u>90210</u> Phone: <u>(555) 123-4567</u> E-mail: <u>j.smith@anytown.com</u> | Date/Time: <u>03/04/2004 10:00 AM</u> Location: <u>123 Main St., Apt. 456</u> City: <u>Anytown, CA 90210</u> State: <u>CA</u> Zip: <u>90210</u> Phone: <u>(555) 123-4567</u> E-mail: <u>j.smith@anytown.com</u> | |
| Vehicle: <u>2000 Ford Focus, Silver, 4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> Mileage: <u>150,000</u> VIN: <u>1F3501A0000000000000000000000000</u> Description: <u>4-Door, 1.8L, 150,000 Miles</u> License: <u>ABC123</u> Title: <u>Clear</u> Make: <u>Ford</u> Model: <u>Focus</u> Year: <u>2000</u> Color: <u>Silver</u> | | | |

[illegible][illegible][illegible][illegible]



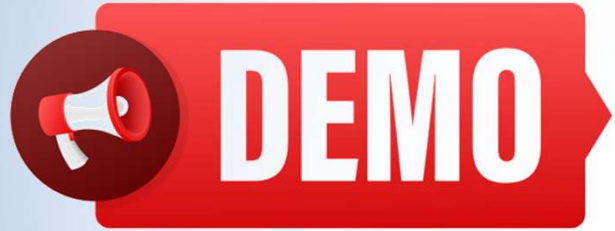
After Scan Correction

Automatic Document Separation Saves Time & Accelerates Capture Workflow



Automatically separate and name multi-page files & storage file folders to accelerate capture workflow.





Form Separation Demo

Poll Question: How does your company or organization add metadata to your scanned documents?

Manually entering the data (e.g. Key from image)

No metadata, we're just creating searchable PDF files

Using automation (Zonal OCR, IDP, RPA, etc.)

We're not scanning documents





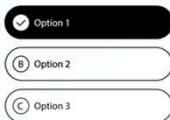
Key From
Image



Advanced
Zonal OCR



Handprint
Field OCR



Choice
List



Checkmark



1D/2D
Barcode



JSON
Parsing



Form
ID



Database
Autofill



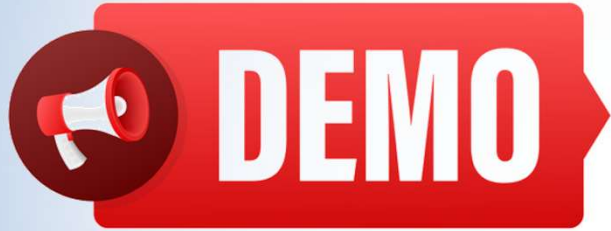
System
Data

The screenshot shows a form titled 'Form A' with various fields and a table of data. The form includes sections for 'Patient Information', 'Pharmacy Information', and 'Prescription Information'. A table on the right side of the form displays data for 'SARAH JOHNSON' and 'Premier Drug Testing'. The table has columns for 'Handprint Text', 'Machine Text', 'Checkbox', and 'Barcode'.

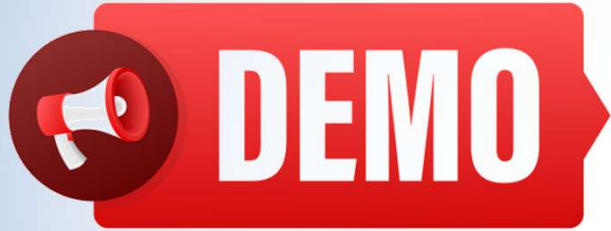
| Handprint Text | Machine Text | Checkbox | Barcode |
|----------------|----------------------|---|-----------|
| SARAH JOHNSON | Premier Drug Testing | <input type="checkbox"/> YES <input checked="" type="checkbox"/> NO | 123456789 |

*“Hand-printed
forms are still a
thing”*





Handprint Recognition Demo

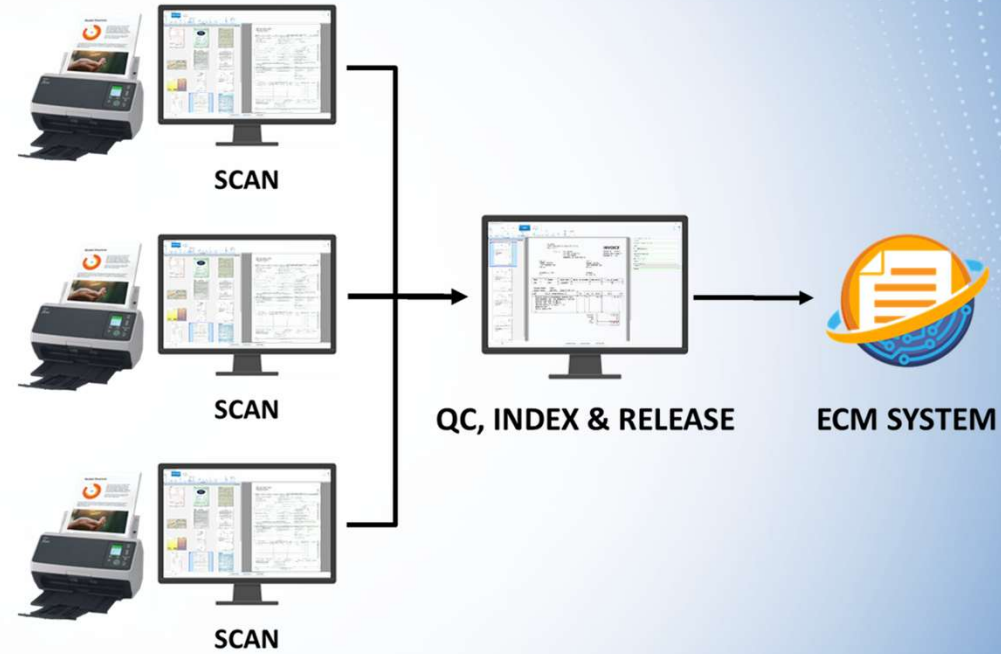
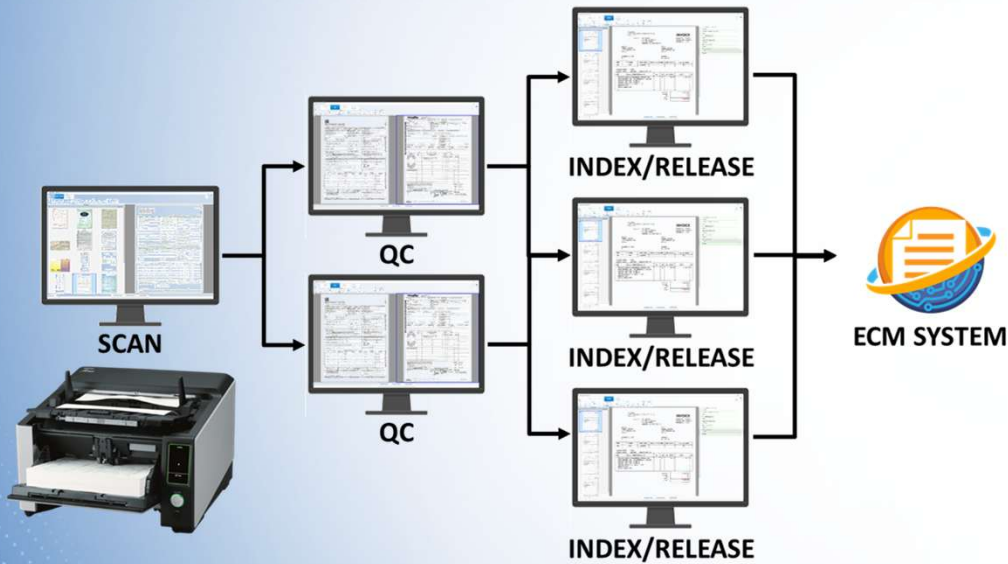


Multi-Form Processing

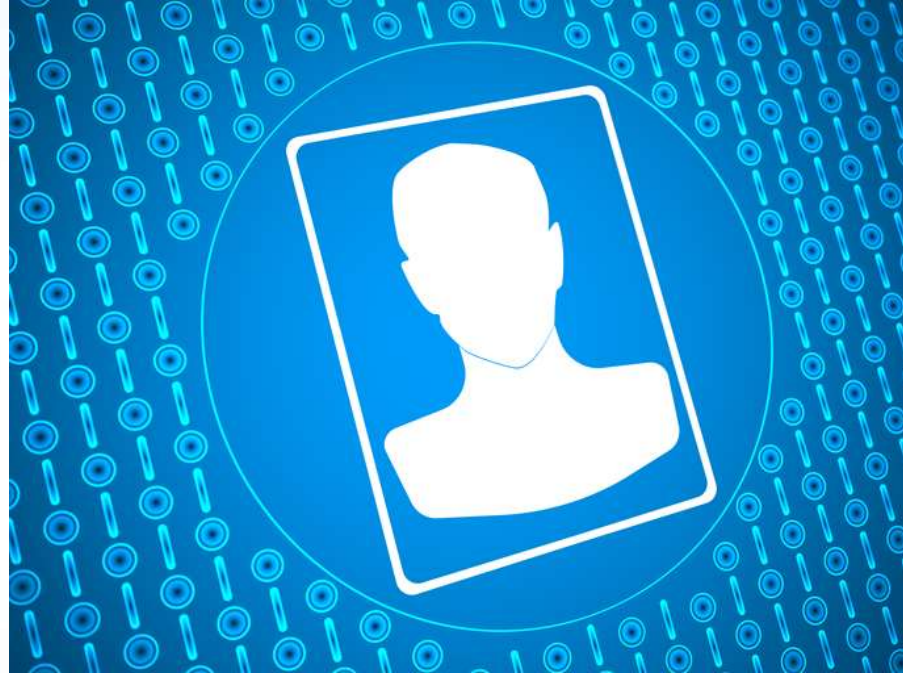
*“We need to
distribute our
capture
workflow”*

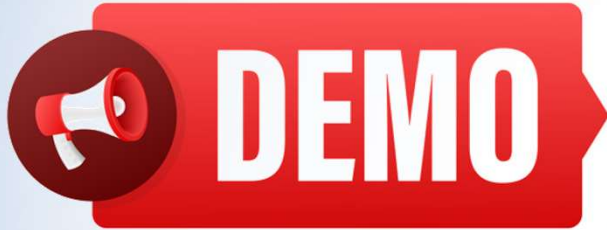


Multi-Station Processing Accelerates Capture Workflow!

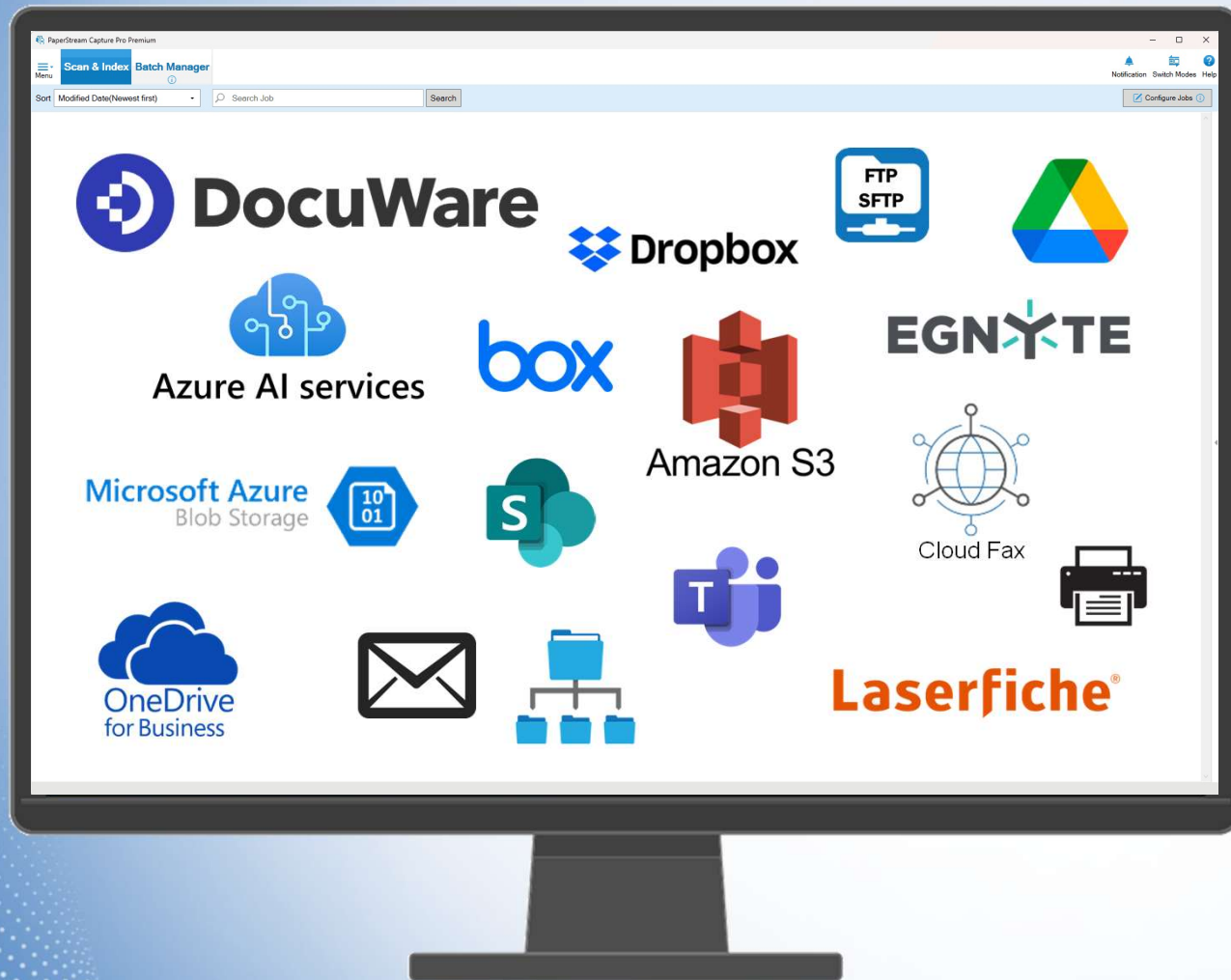


*“We need to
protect
confidential
information”*





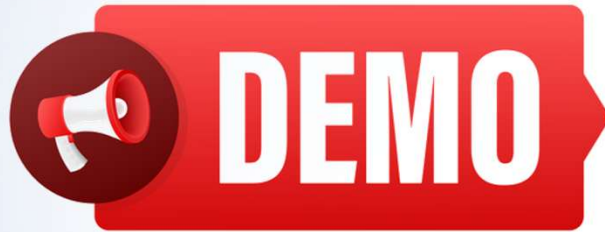
Field Redaction



**PaperStream
Capture Pro**



**Direct Application
Integrations Streamline
Capture Workflow &
Reduce Manual Processing**



Application Integration

PaperStream

Capture **Pro**

Contact us to learn more!



**Come by our booth to see our scanning
& software solutions in action**




GLOBAL SUMMIT





PS PaperStream
Capture **Pro**

Thank you!