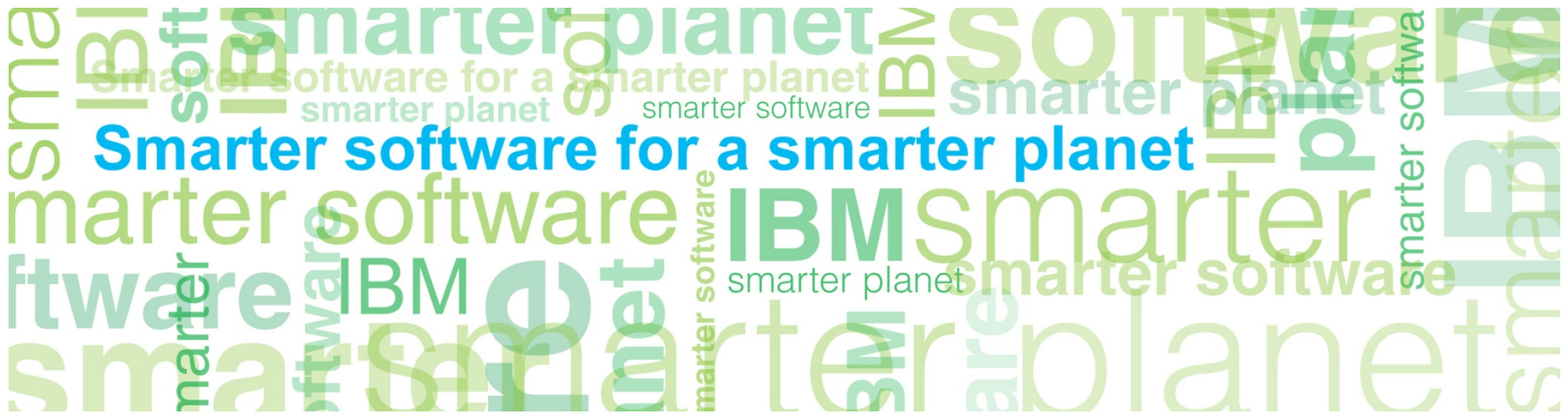


# The Five Phases Of Capture

Presenters: David Jenness, Kevin Craine

May 12, 2011



## Your Presenters

### Kevin Craine

- Author, "*Designing a Document Strategy*," sold in over 30 countries and host of the Document Strategy Podcast
- 25 years managing large document services organizations in the health care, pharmaceutical and aerospace industries
- Founding editor of "*Document*" magazine
- Managing director of Craine Communications Group



### David Jenness

- Market Segment Manager, Document Imaging and Capture, IBM Software Solutions Group



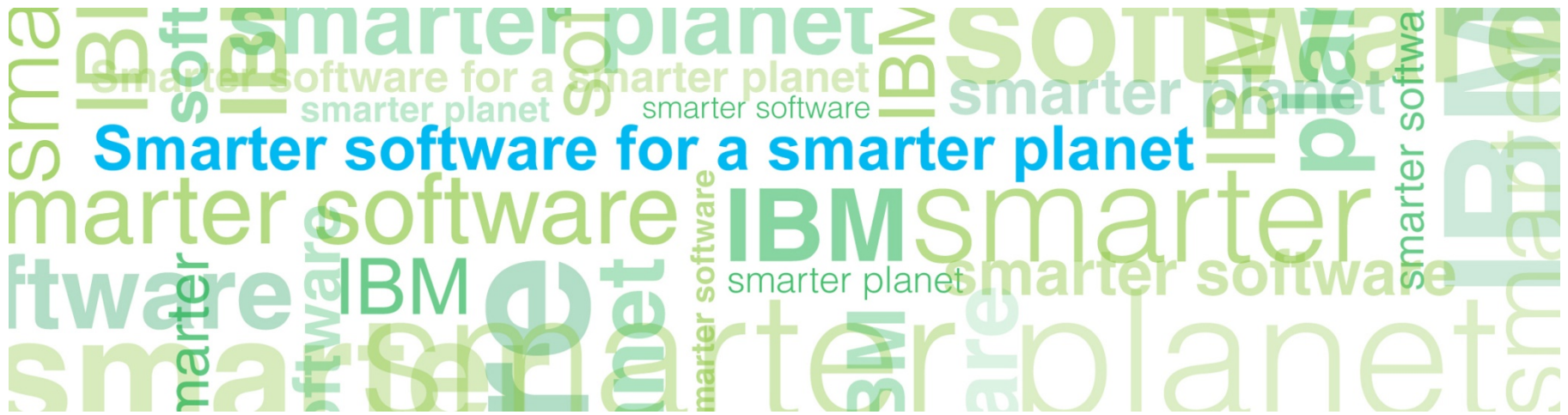
## Agenda

- The Five Phases of Capture
  - What are the 5 Phases?
  - What are the benefits of each?
  - How do you move from one to another?
  - Use cases
- About IBM Datacap Taskmaster Capture
- About IBM Production Imaging Edition
- Q & A



---

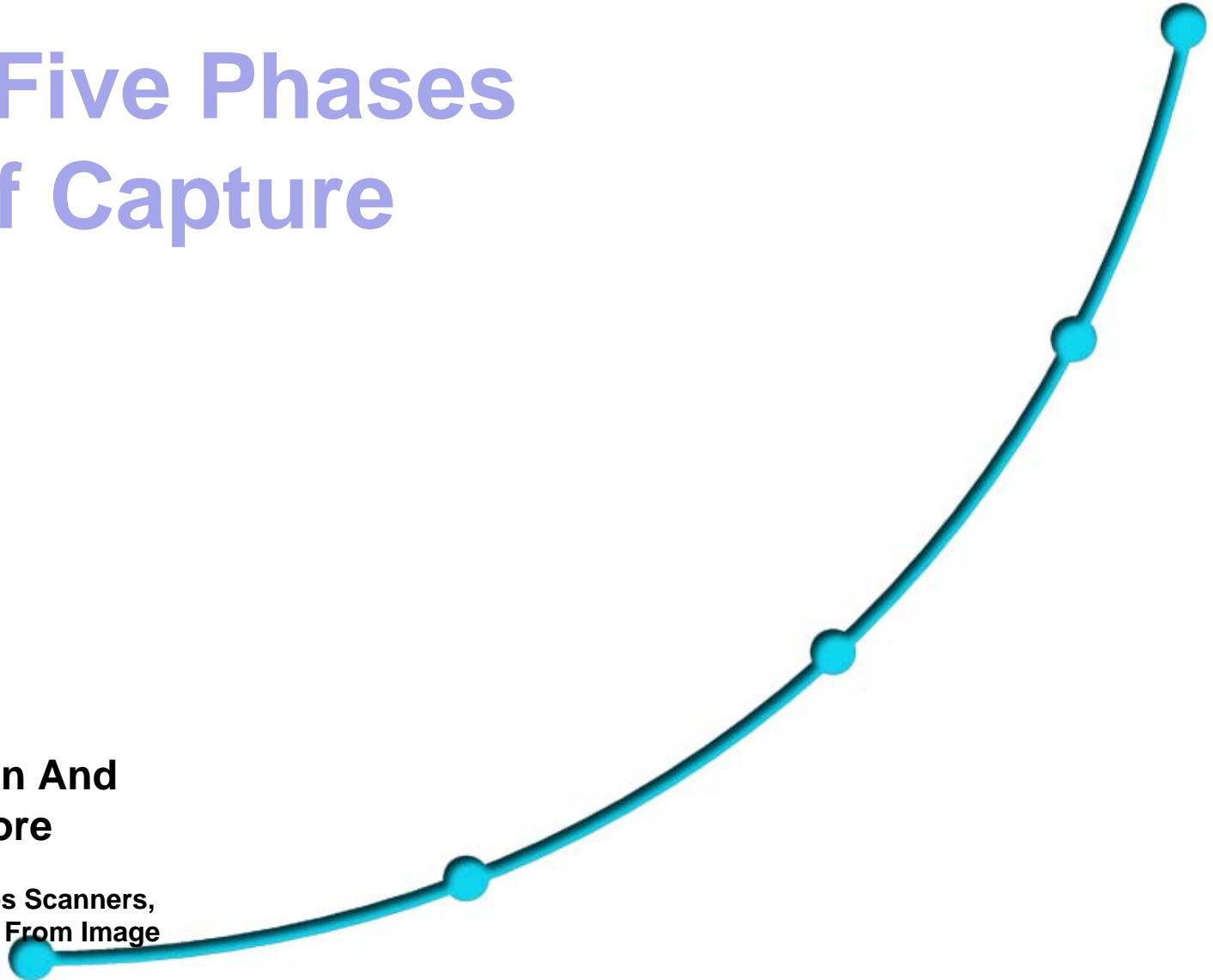
# Document capture means many things to many people



# The Five Phases Of Capture

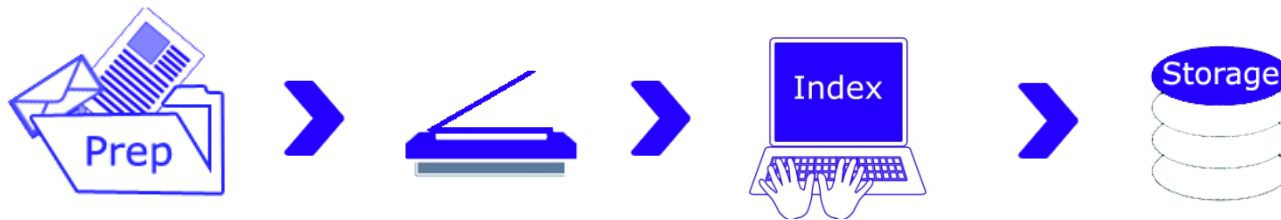
## 1. Scan And Store

Capture drives Scanners,  
provides Key From Image



# Scan and Store

- ✓ The First Phase of Capture involves presorting documents, scanning and adding indexes manually.
- ✓ Nearly every ECM software solution offers this type of capture with their ECM platform – many ship standard with ECM.
- ✓ Without document automation, each step involves manual labor





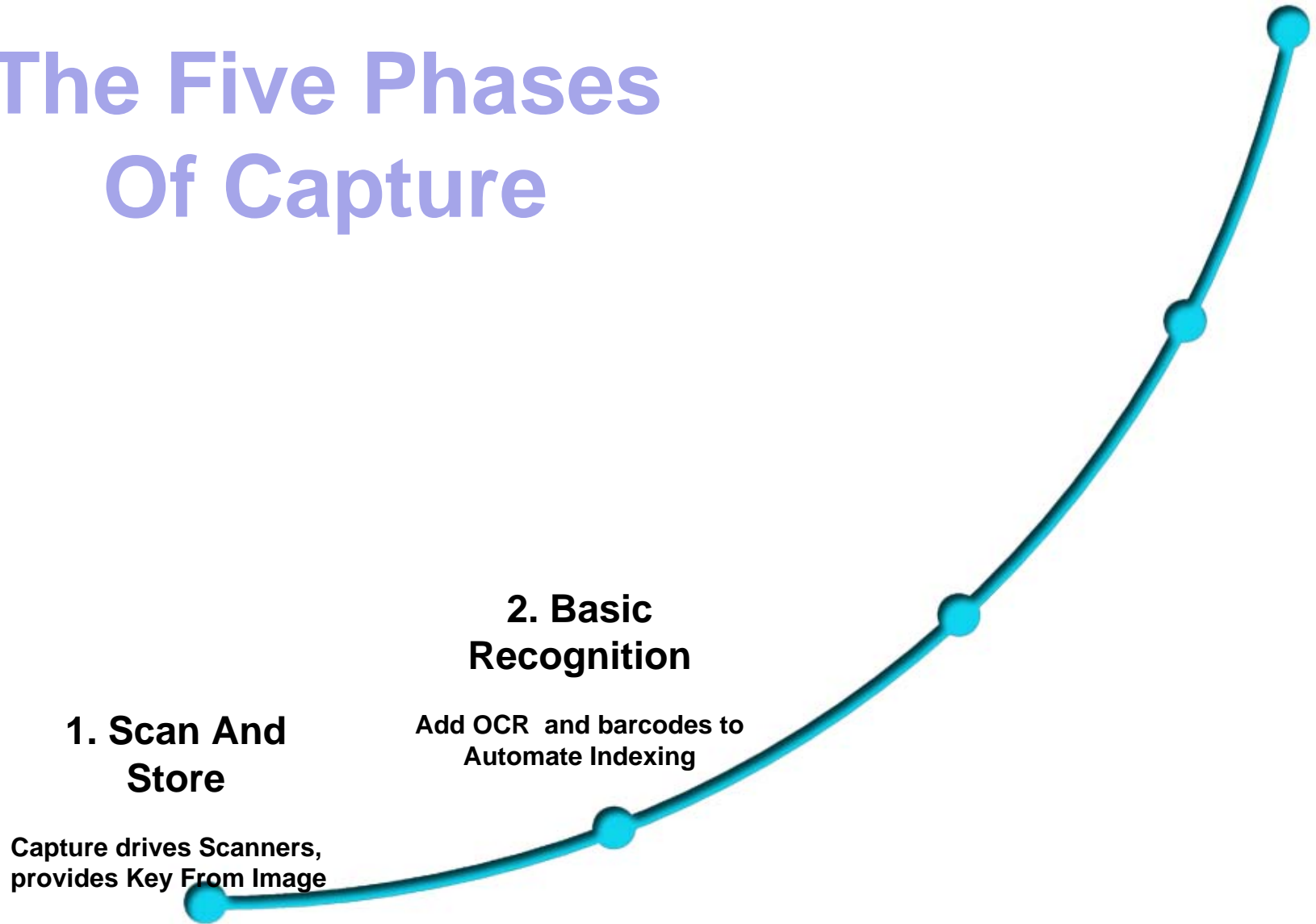
## County Government Scans Healthcare Documents

*"We've gotten rid of 16 file cabinets and two storage spaces and we've still got a long way to go!" –  
Administrative Assistant, County*



- ✓ Scanning health and legal records going back 100 years to free up storage space.
- ✓ A Phase One capture system drives the scanners and presents each operator with several document types to choose from.
- ✓ Once the operator selects the document type from a dropdown menu, several data fields appear, which are manually entered for index purposes.
- ✓ The image is routed to an image repository and the paper can be disposed.

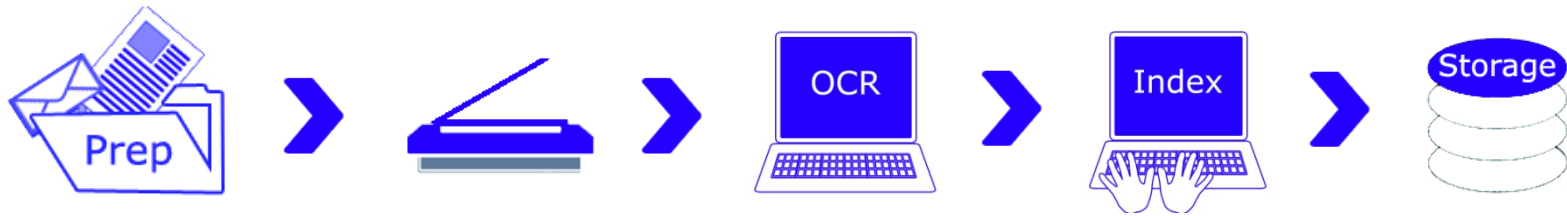
# The Five Phases Of Capture





## Basic Recognition

- ✓ The Second Phase of Capture involves prepping documents and using basic recognition of barcodes, machine printed text or check boxes on separator sheets and document header information.
- ✓ The recognition is used to create the meta-data for indexing document images for efficient storage and retrieval in an ECM repository.



## Manufacturing Company Image-enables Human Resources

*"I used to spend several hours scanning and typing data, now I just scan." – Human Resources Manager*



- A small manufacturing company is using a basic capture system to scan Human Resources documents, including applications, payroll, benefits, employee review and correspondence.
- Before scanning, each document type is separated and a document ID sheet with bar code and checkboxes is placed in front of it.
- When scanned, capture “reads” the separator sheet and applies the proper document ID and index data to the document that follows.
- Index data is checked by a verifier and the image is routed to storage

# Recognition-Assisted Indexing



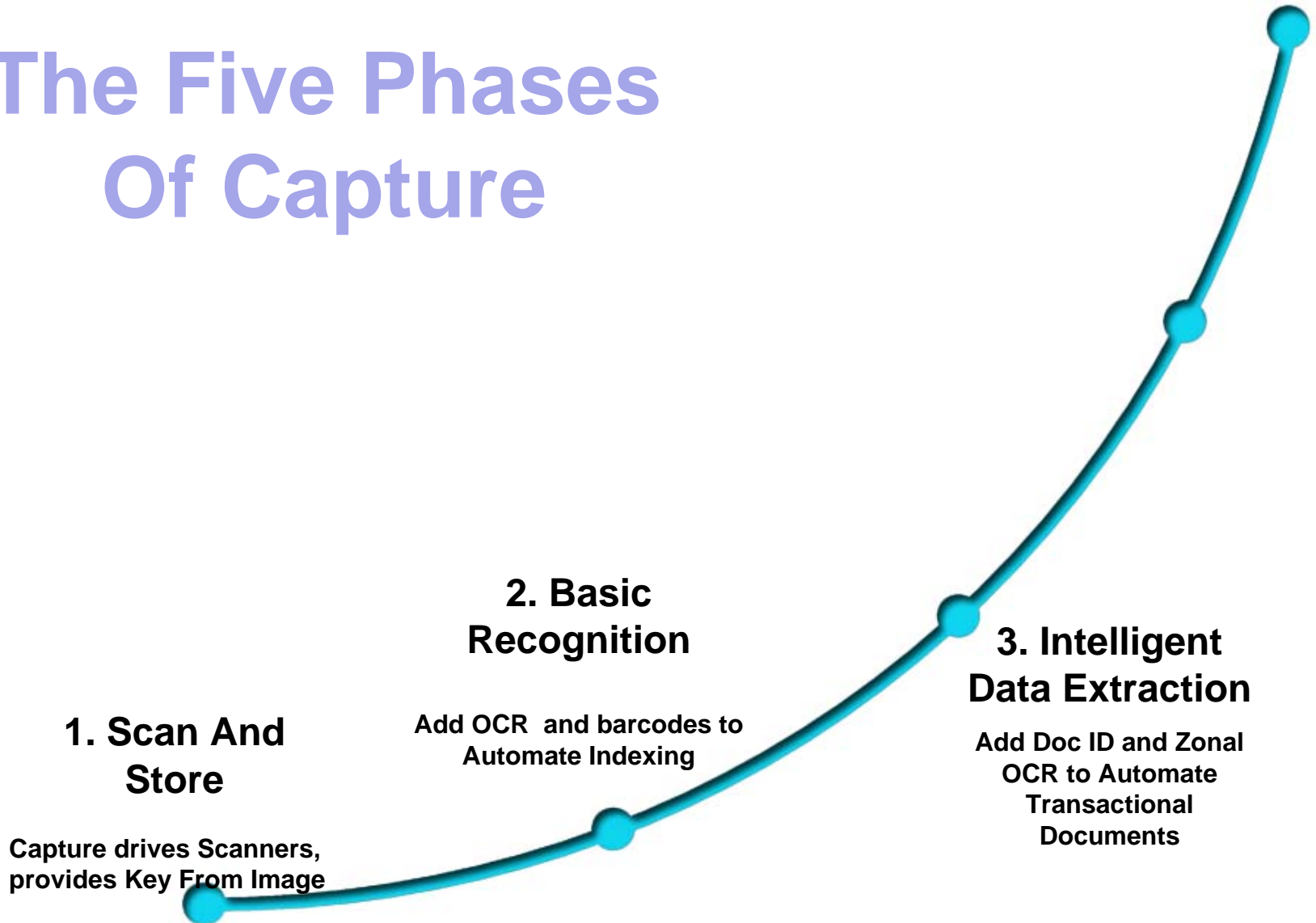
## Document Separator



**Please select the valid policy “Bucket” (choose only one per Separator page):**

1. Original Policy .....
2. Post-Issuance Endorsements .....
3. Post Binding Correspondence .....
4. Coding Information .....
5. Binder .....
6. Pre-Binding Correspondence .....
7. Surveys/Inspection Report .....
8. Certificate of Insurance .....
9. Premium Finance Agreement .....

# The Five Phases Of Capture



## Intelligent Data Extraction

- ✓ The Third Phase of Capture involves prepping the documents and using advanced recognition to locate and extract typed and hand printed data from forms and “transactional” documents, such as invoices, claims, tax returns, and applications, etc.
- ✓ Intelligence comes from capture system’s ability to manage variability with flexible document identification and data location methods as well as multiple recognition strategies.





## Intelligent Data Extraction Requires The Ability to Handle Variability

- ✓ Multiple strategies for identifying document types automatically
  - ✓ Pattern matching
  - ✓ Key word search
  - ✓ Image matching
  - ✓ Text analytics
- ✓ Multiple strategies for locating data on variable documents
  - ✓ Location rules
  - ✓ Pre-set zones
- ✓ Multiple recognition engines to extract:
  - ✓ Hand print
  - ✓ Machine print
  - ✓ Bar code and 2D bar code
  - ✓ Check box
  - ✓ Dual engines with voting

# Dual Recognition Engines and Raise Accuracy by "voting"

Engine A

Engine B

789012345VP6789

789012345VP6789

999699997887999

999599996998499



789012345VP6789

9999999991999999

# Rich verify environment

Database lookups

Full Document Image

Image and data snippets

Line Item Snippets

Line Items in spreadsheet format

The screenshot shows a software application window titled "Verify - Batch Pilot". It features a main form on the left for data entry and a large "Image View" window on the right displaying a scanned invoice. A smaller spreadsheet window titled "APT Detail dit" is also visible, showing a table of line items.

**Main Form Data:**

Remittance_Zip	80202-4628
Vendor_Number	TS13687
Invoice_Type	
Invoice_Date	02/26/07
Invoice_Number	# 28100
Tax	77.39
Invoice_Total	1551.39
PO_Number	
Shipping	

**Invoice Image Content:**

**JWS of Colorado, Inc.**  
 1675 Broadway, Suite 2800  
 Denver, CO 80202-4628  
 Phone (303) 825-5467 Fax (303) 825-3434

**HC technologies**  
 4903 Business Parkway  
 Rochester, MN 55901  
 (507) 555-1212

**INVOICE**

salesperson	job	SHIPPING METHOD	shipping terms	delivery date	PAYMENT TERMS	due date
195				2/10Net30	03/26/07	

**Line Item Table:**

qty	Item #	description	unit price	discount	line total
10	201	Trucking - Ticket 138432	70.00		700.00
130	204	Brine Water	.70		91.00
130	203	Fresh Water	.25		32.50
19	206	KCL	6.00		114.00
6	201	Trucking - Ticket 146611	65.00		390.00
130	203	Fresh Water	.25		32.50
19	206	KCL	6.00		114.00

**APT Detail dit Spreadsheet:**

Line Number	ItemID	ItemDesc	Qty	Price	LineTotal
1	201	Trucking Ticket 138432	10	70.00	700.00
2	204	Brine Water	130	0.70	91.00
3	203	Fresh Water	130	0.25	32.50
4	206	KCL	19	6.00	114.00
5	201	Trucking Ticket 146611	6	65.00	390.00
6	203	Fresh Water	130	0.25	32.50
7	206	KCL	19	6.00	114.00

**Summary Totals:**

Tax	77.39	SubTotal	1551.39
Shipping	0.00	Invoice Total	1551.39
LineItem Total	1474	Difference	0.00

# Dow Jones & Company Automates Subscription Processing

“Advanced Capture enables us to scan and capture critical customer information and quickly process our subscription orders.”

—Brian Gilmore, national manager of circulation service operations for Dow Jones

- Dow Jones’ central processing center receives 8,000 subscription postcards a day – up to 50 different offers, each with its own layout.
- Each card has a unique 11 digit code beginning with the letter B.
- Capture searches the document for the code, using intelligent location rules. With the number, Taskmaster retrieves the offer details associated with each card.
- Capture extracts name and address and length of subscription information and passes it for processing.
- Detailed reporting informs Dow Jones of its cash position at the end of each day.



**What's News—**  
Business and Finance

**Johnson and Johnson** were part of a buyout...  
**Shares in a bid** to sell part of all of the 98...  
**The SEC is to decide** whether to weigh in on an...  
**Private equity** firms...  
**Some private equity** firms are...  
**IBM agreed to buy** Intel...  
**China's economic**...  
**China's economic**...

**Playing Cellphones On Stage Has Ring Of Respectability**  
Brian Yoon's Music Career Takes a High-Tech Turn: Listening for a Special Plug  
By Anna Hazzard

**At Morgan Stanley, A Game of Catch-Up**  
To Regain Firm's Peak, CEO Must Pushes Big Bets, Big Deals  
By Anna Hazzard

**COMEBACK PLAN**

**Return Engagement**

**Florida Tackles Task of Cutting Property Taxes**



Anna Hazzard



John J. Mori

## Departments Where Capture Makes A Difference

*Which of your enterprise systems are capture-enabled and integrated at a process level?*

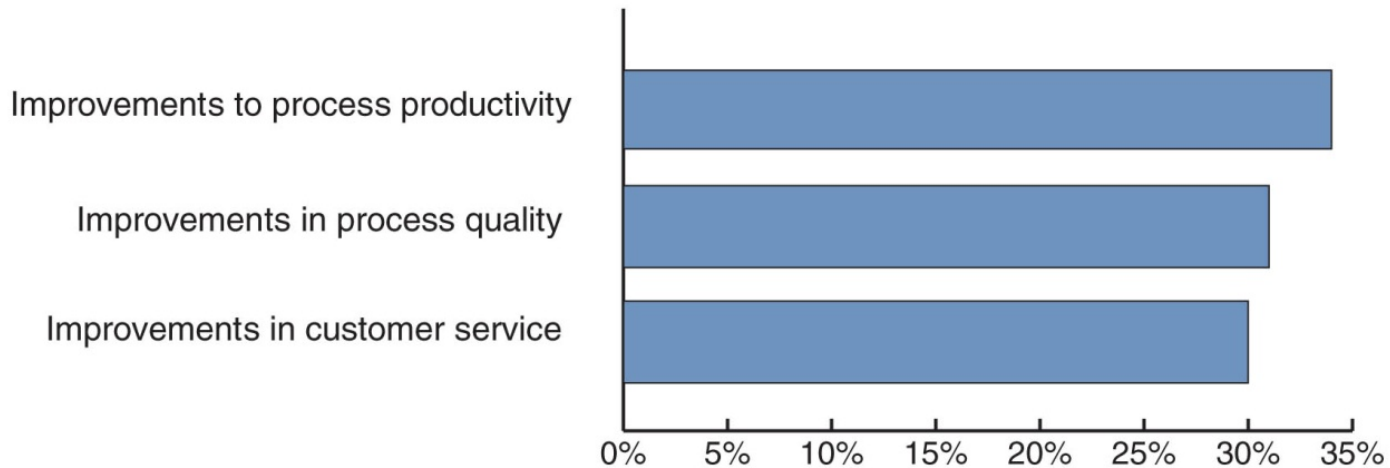


Capture and Business Process. AIM Market Intelligence. 2010



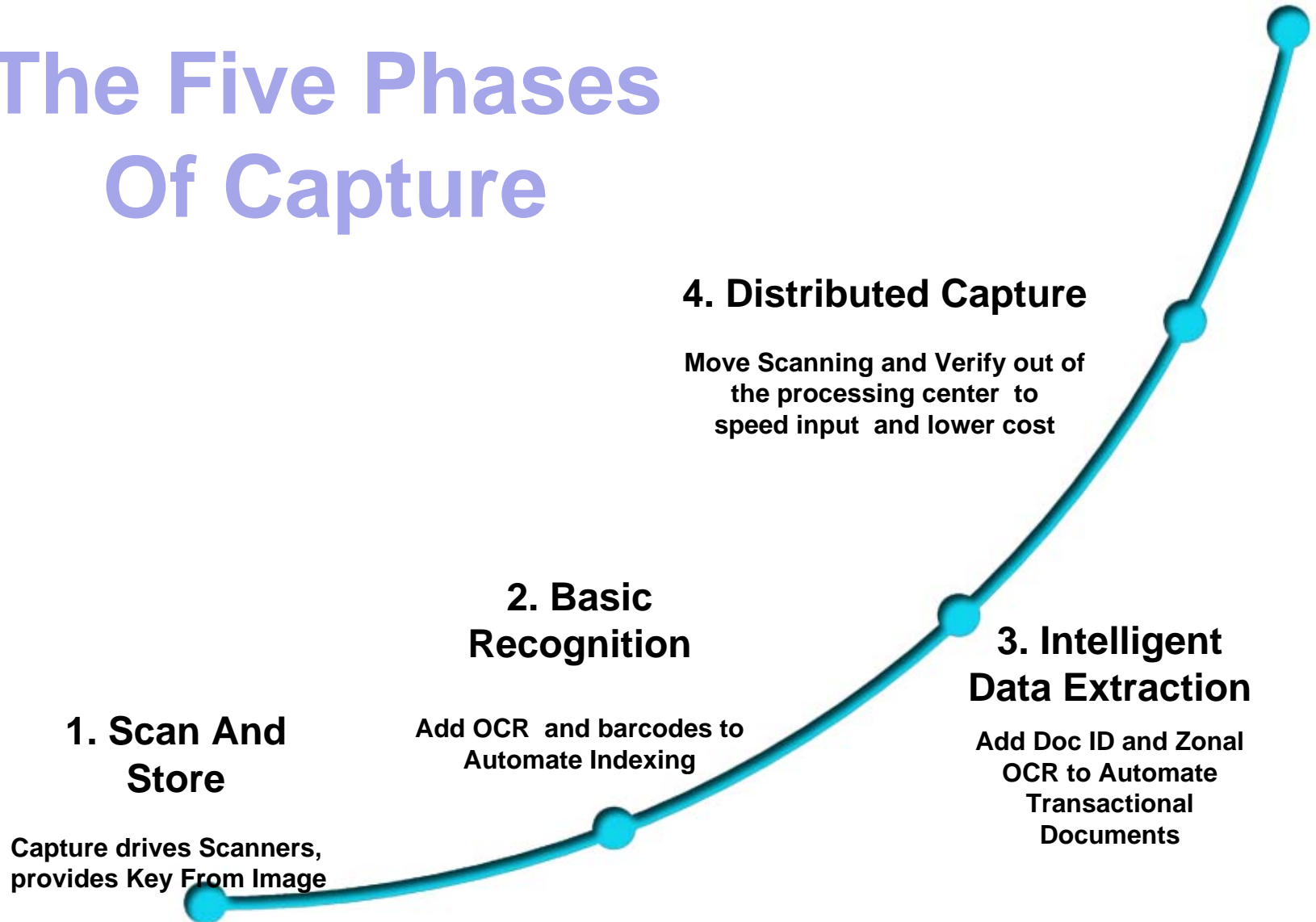
# Intelligent Data Extraction

*Which aspects of data recognition provided the best financial return on your investment?*

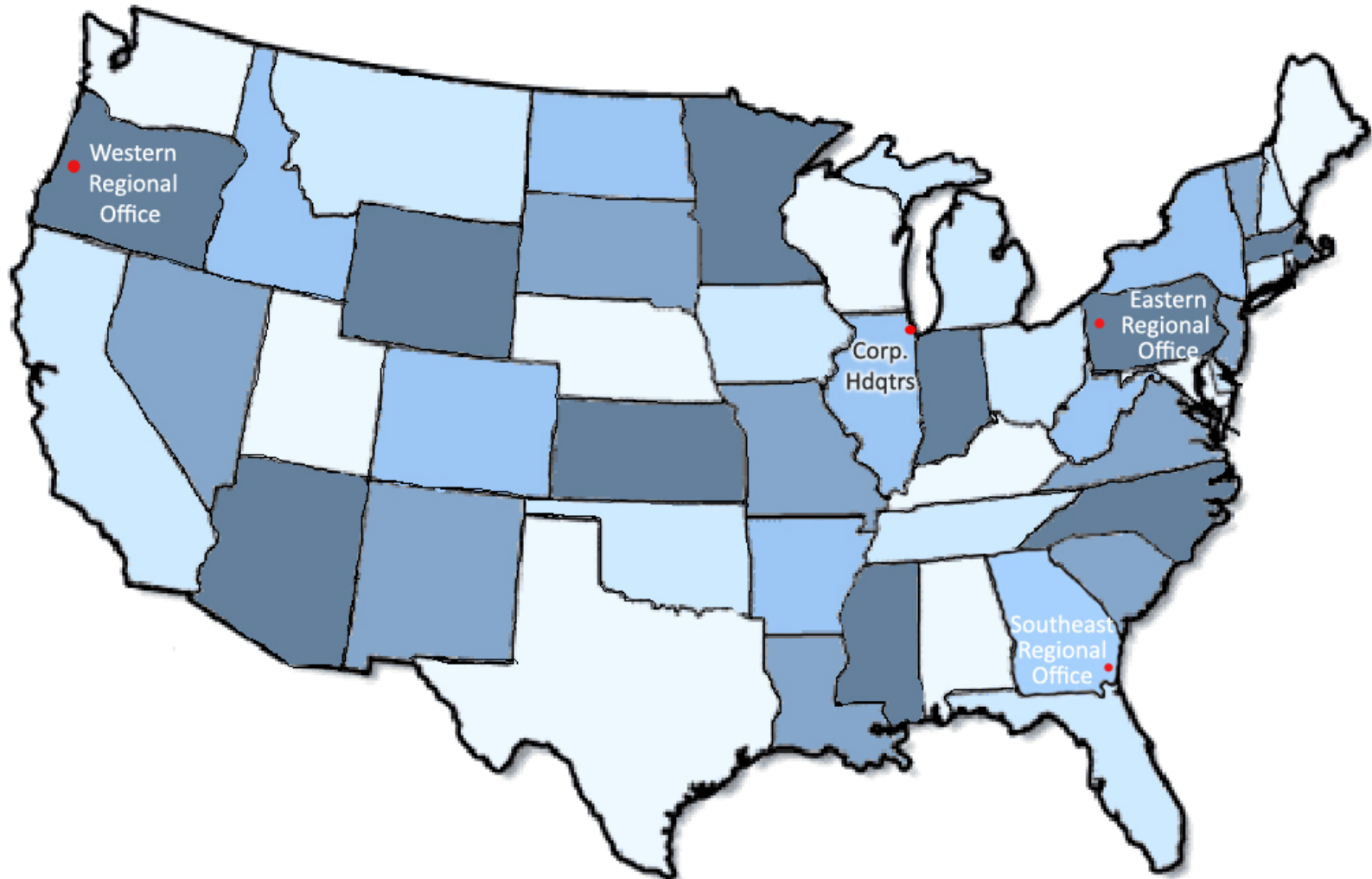


Capture and Business Process. AIM Market Intelligence. 2010

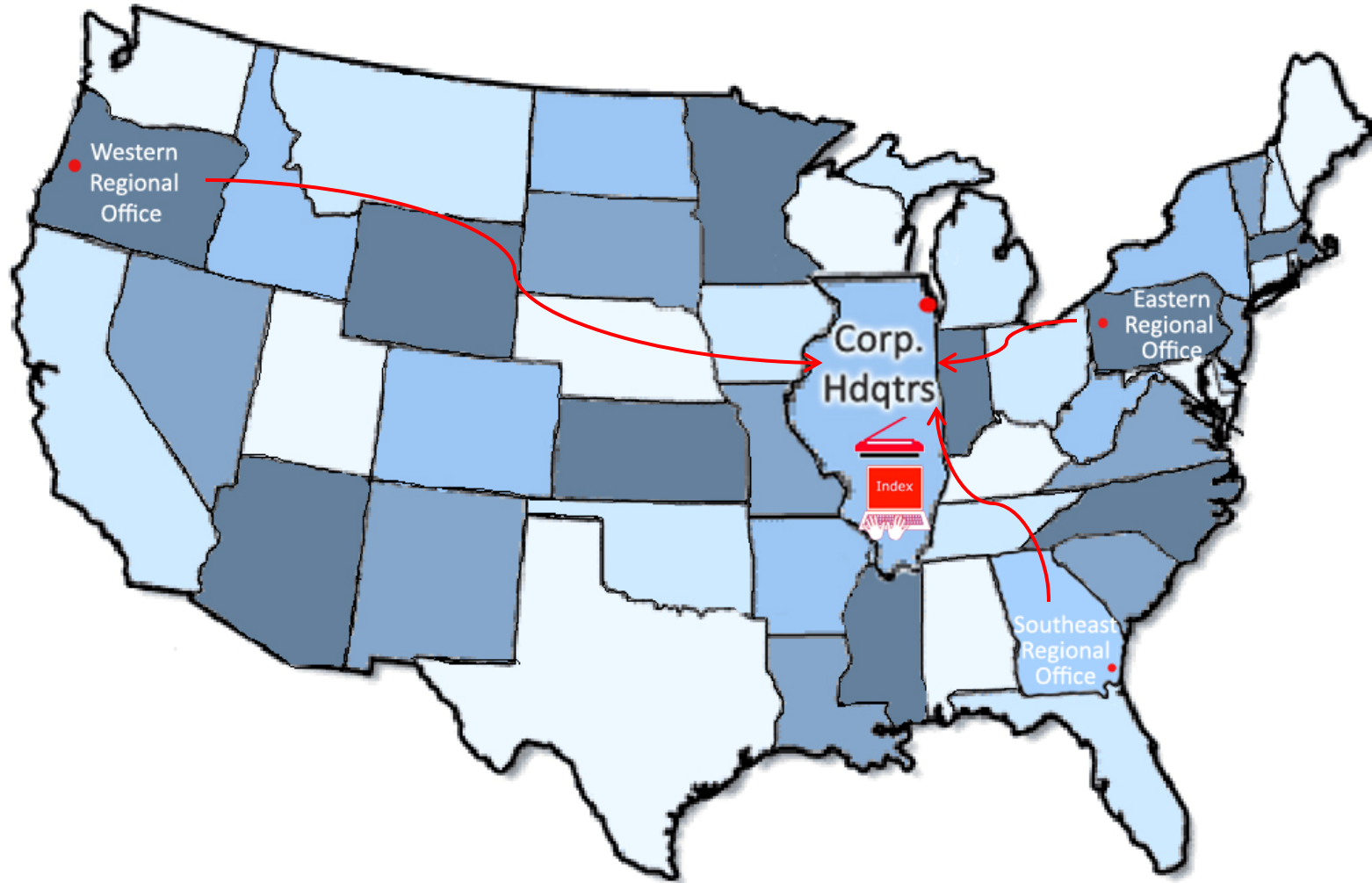
# The Five Phases Of Capture



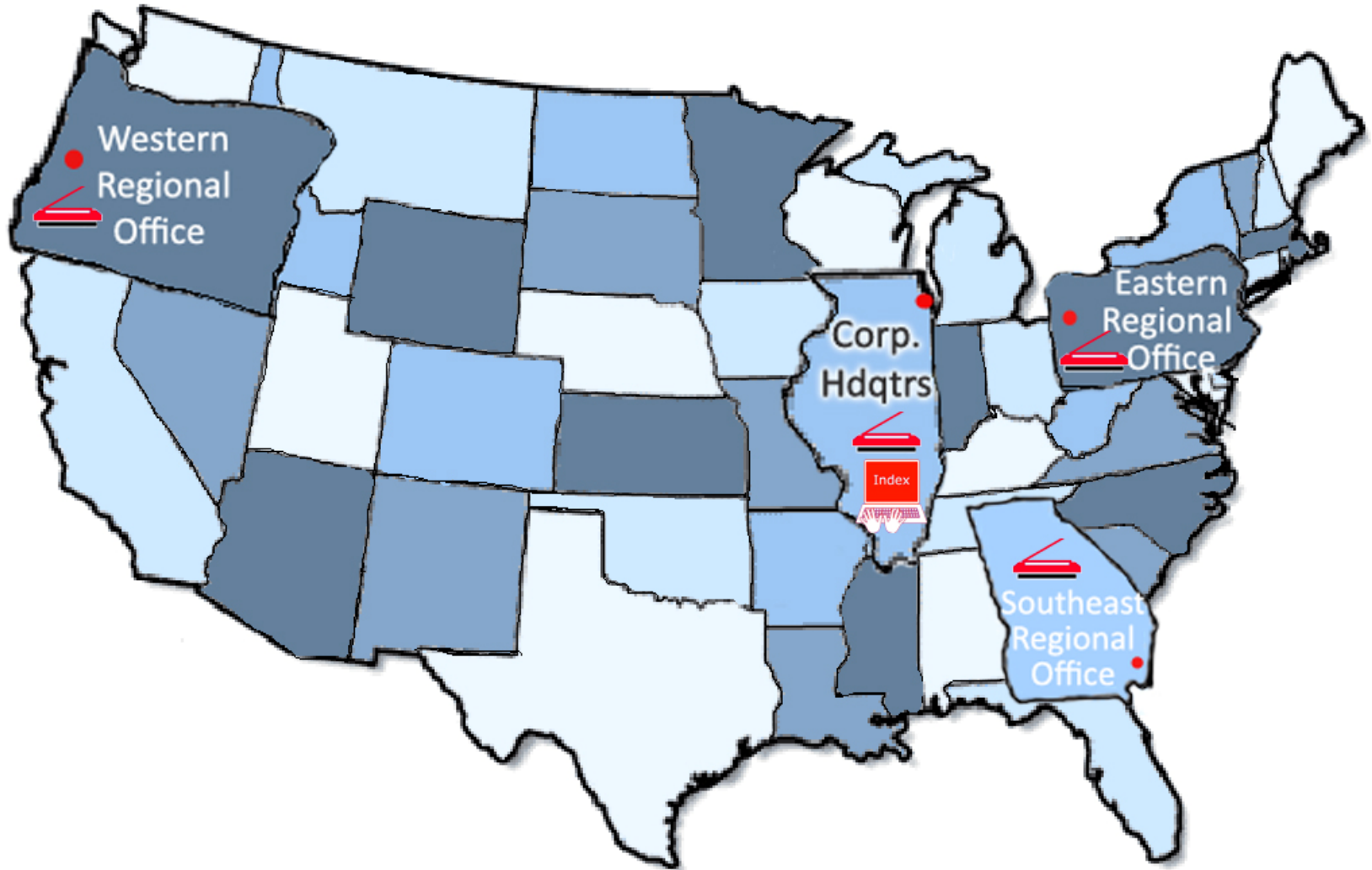
## Distributed Capture Scenario: A company has four offices in North America...



# 1. Organization ships documents to headquarters for scanning in the central scan center

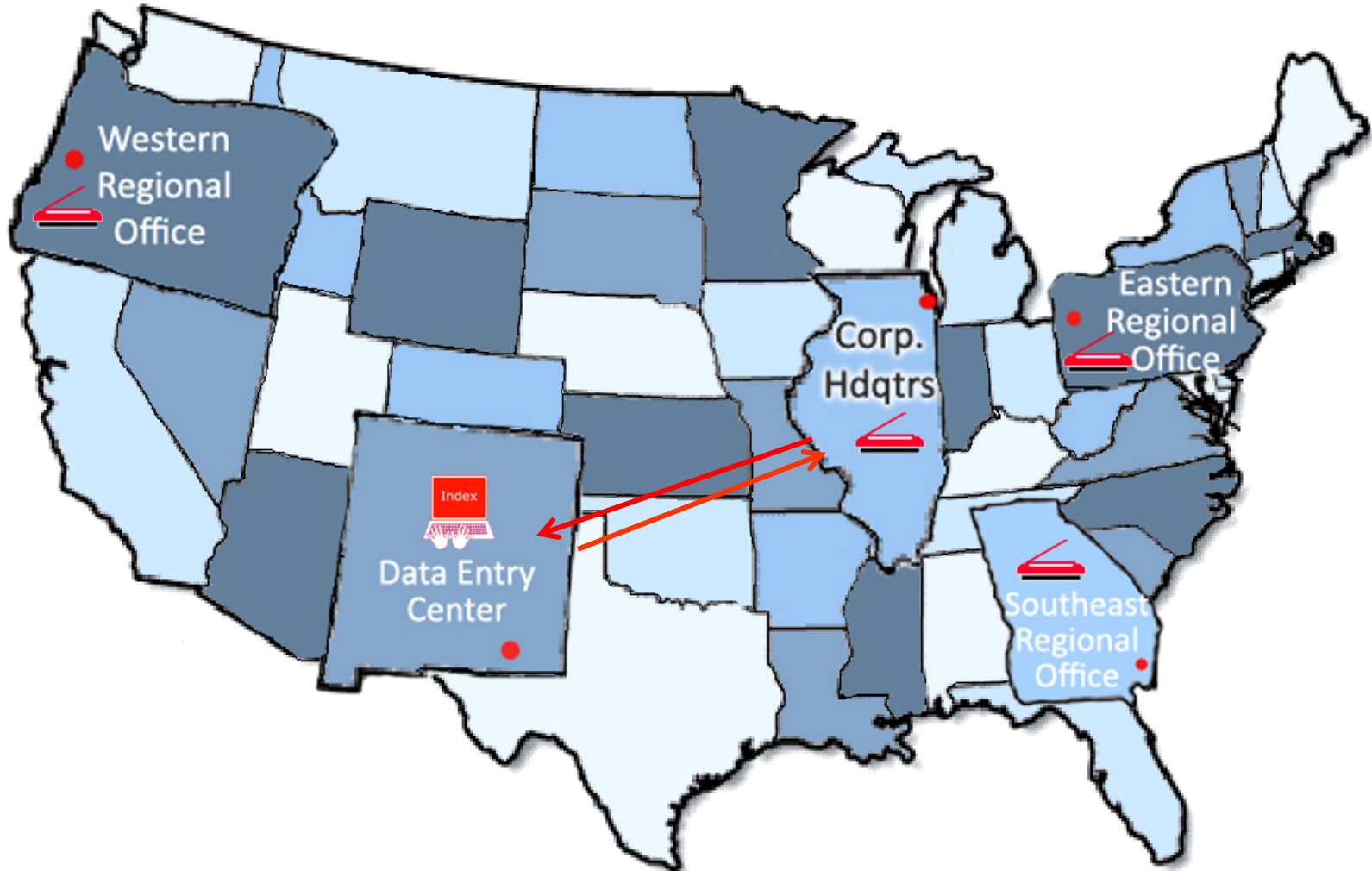


2. By scanning at the regional centers and sending images over the internet to headquarters, they eliminate document shipping costs and time





### 3. By moving verify and indexing tasks to lower cost regional center, they save additional cost



# Distributed Capture Benefits

- Cost Savings
  - Postage, overnight mail, presorting
  - Documents lost in transit
  - Move work to regions with lower wage rate
- Accelerate Processing Time
  - Immediate scanning and input
- Increase Accuracy
  - People who generate documents have better knowledge
- Project Management
  - Overflow and spikes in volume
  - Migration to new processes

## Murphy-Hoffman Trucking Company Eliminates Shipping Costs

*“Now staff at headquarters isn’t waiting until paper arrives to perform their work. They always have work available.” – Imaging Manager, Miurphy-Hoffman Trucking Company*



- 65 regional sales and service centers throughout the Midwest
- Replaced overnight shipping expense with 65 scanners and browser-based scanning software
- Invoices, sales, lease and service documents are scanned as soon as they are generated
- Uploaded to Kansas City headquarters for processing and storage
- Now documents are available immediately
- Staff at headquarters no longer wait for documents to arrive
- Document shipping expense eliminated – MHC estimates that they save \$442,000 per year

## Virginia Department of Taxation enables workers in low income areas

*"We realized we could use the thin client to have at-home workers do data entry and verification of returns."*

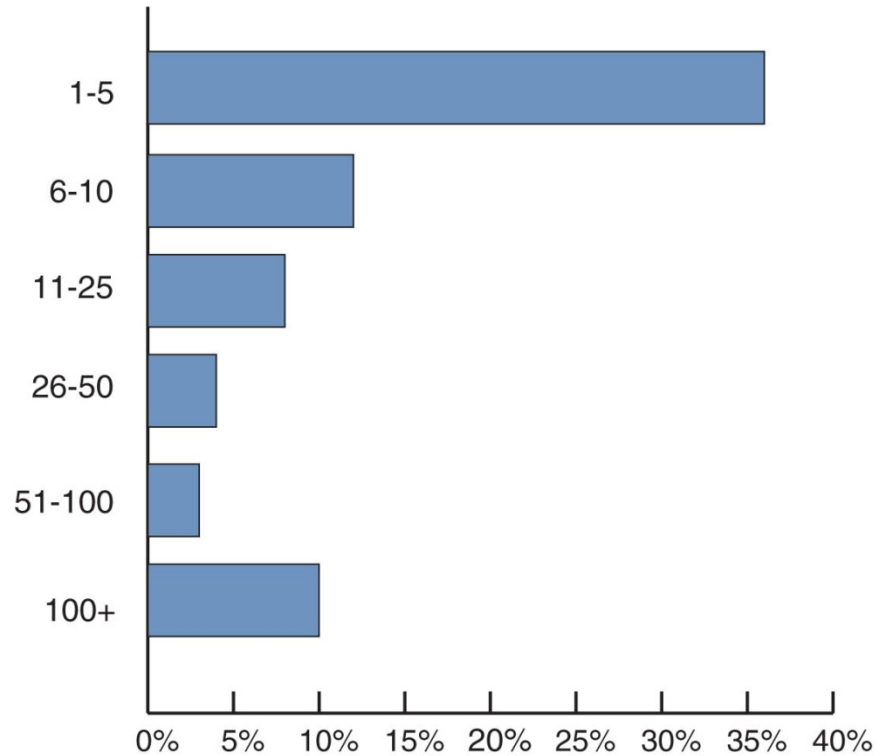
*— Nancy Wilson, Virginia Tax's Manager of Automated Processing Systems*

- Processing 1.5 million paper tax returns every year, scanned in Richmond processing center
- Each captured return is presented to a verify operator who confirms data accuracy and fixes low confidence characters when needed
- Virginia passed a law in 2008 to stimulate jobs in low income areas
- Virginia Tax distributes a percentage of tax returns to At-home workers in low income areas using browser-based verify panel
- Distributed capture helps Virginia deliver on its economic pledge and provides Virginia tax processing staff with maximum flexibility



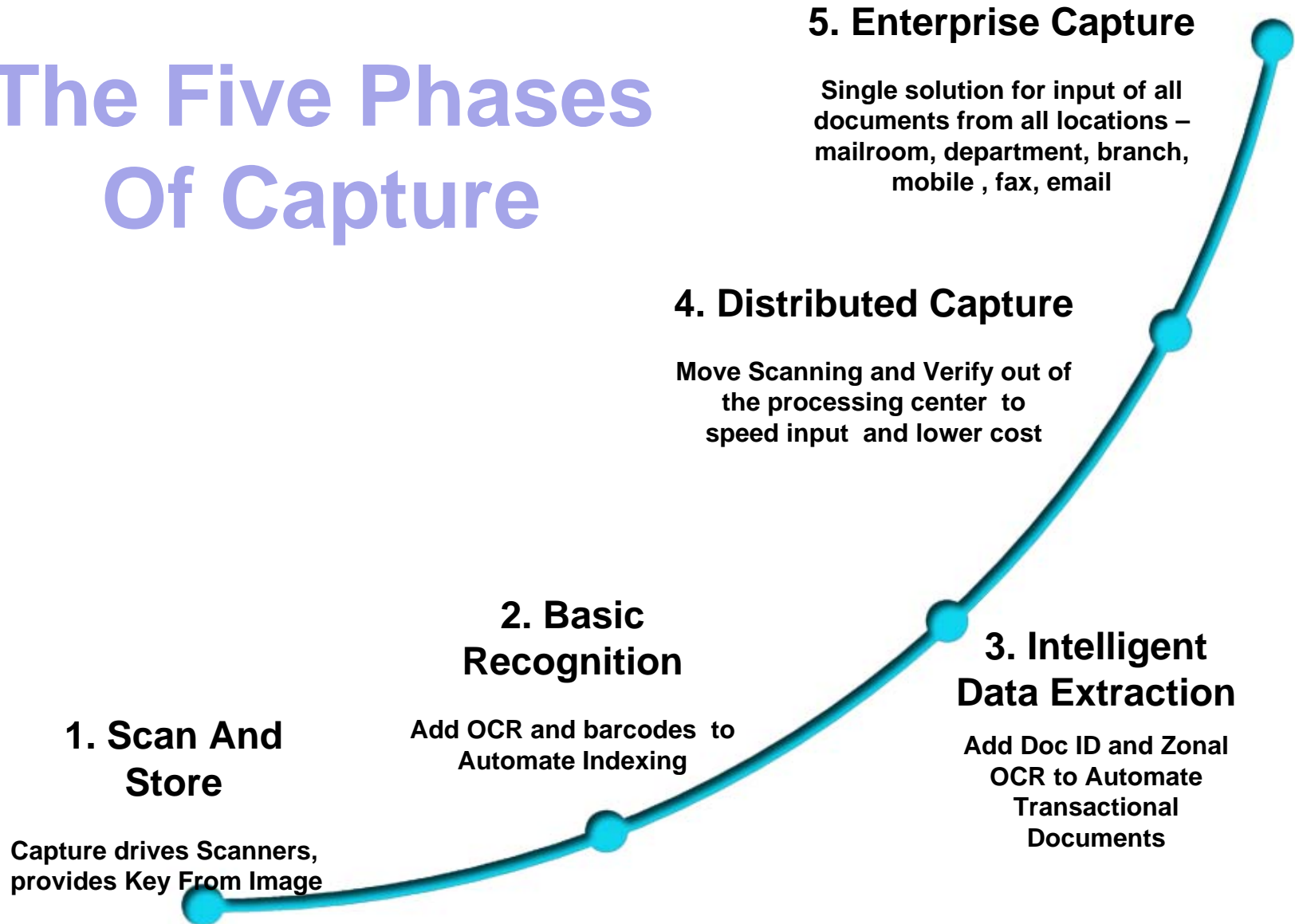
## Distributed Capture Grows In Popularity

*How many distributed capture sites do you have across your enterprise?*



Distributed Capture Survey Report 2009. Datacap.

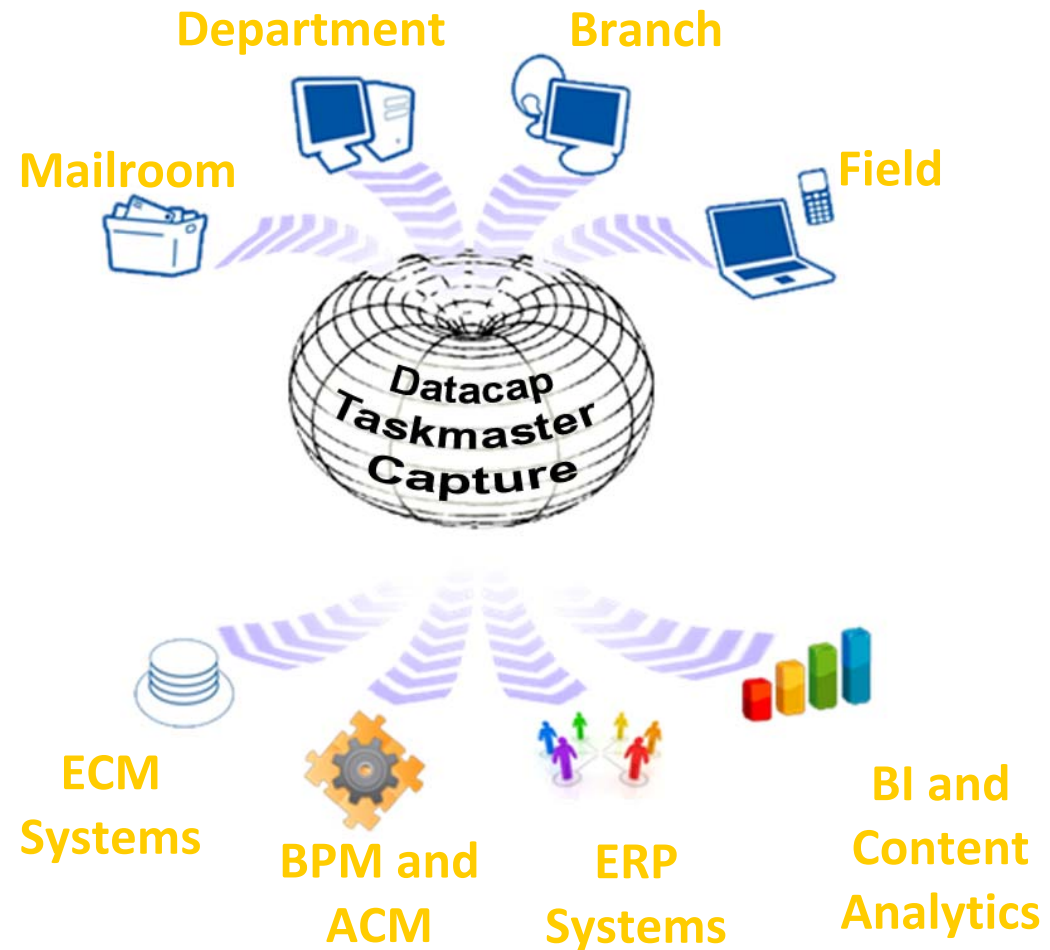
# The Five Phases Of Capture





# Enterprise Capture

- ✓ A universal capture portal that can transform all documents
- ✓ Capture documents at every entry point in the Enterprise
- ✓ Input any mode for consistent processing rules
- ✓ Point and click capture process management enables clients to orchestrate complex capture solutions – without waiting for expensive programmers to build an application



## BlueCross BlueShield Health Insurer Captures All Documents

*"I can see us adding new documents to our capture portal for a very long time."*

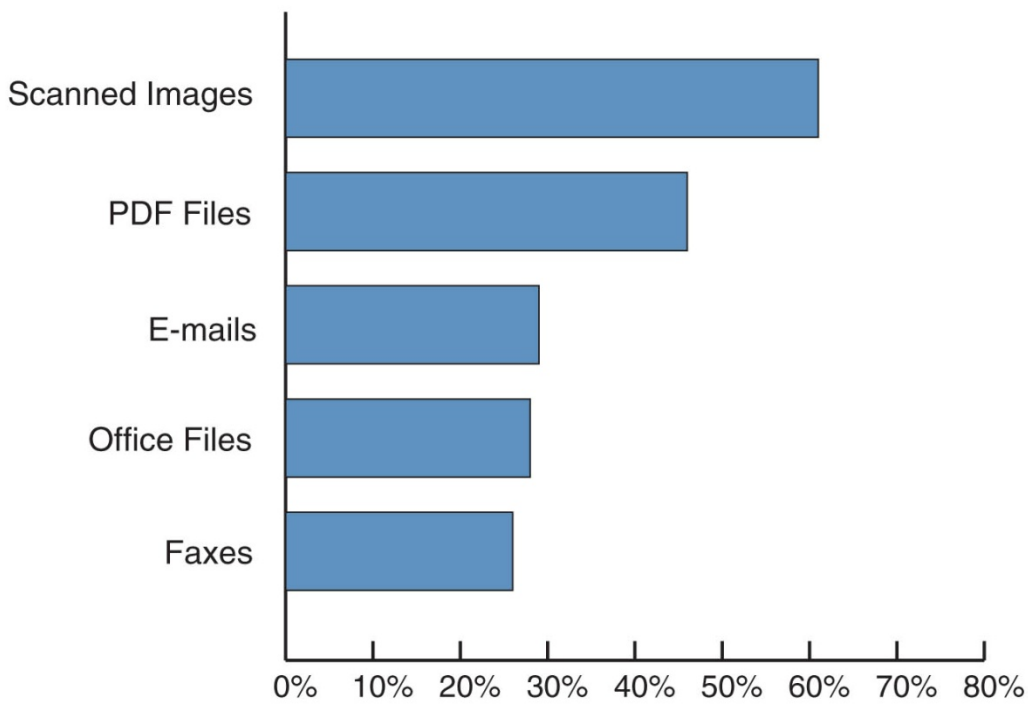
*— Claims Manager, Major BCBS*



- Purchased IBM Datacap Taskmaster Medical Claims Capture to automate input of 12,000 paper health claims a day
- Reduced labor by 50% and shrunk turnaround time
- Made a strategic decision to extend capture to other departments
- Began a process of adding one or two departments a year
- Now they are scanning 50,000 documents a day, including:
  - Contracts
  - Enrollments
  - Medical tests
  - Invoices
  - Human Resources
- Added remote scanning from 5 satellite offices
- Added fax capture
- Email and Electronic documents on horizon

# Need To Input A Mix Of Documents

*Which of the following content types are you processing for automatic capture of data?*



Capture and Business Process. AIIM Market Intelligence. 2010.

# Global Logistics Company improves productivity and service

*Represents the state of the art for capture today: capturing paper, fax and emails, distributed scanning from many different sites, with many rules-driven variations.*



- 150,000 documents arriving every day from every source – mail, fax, email - and piling up rapidly as company prepared customs paperwork for shipments. Customs has many requirements for complete declaration at border crossing
- Deployed seven imaging applications enabling faster order processing with fewer errors
- Process ~600,000 pages per day in U.S. (~3,000 users) and expect to process ~4 million pages per day (~10,000 users) globally.
- Company is able to move more shipments across borders with 30% less resources with reduced lost documents and data errors while also improving cycle times and accuracy.



# IBM Datacap Taskmaster Capture

*The Enterprise Capture Platform*

# The AIIM 2011 Capture Software Product Survey

**April 2011 - IBM Datacap Taskmaster Capture scored the highest total score for Transactional Batch Capture based on 12 categories:**

1. Ease Of Use
2. Ease Of Roll Out/Scalability
3. Ease of Configuration/Customization
4. Security
5. Strategy
6. Stability/Bugs
7. Integration
8. Community and Channel
9. Support/Maintenance
10. Value for Money
11. Standards
12. Application Development

IBM DataCap Taskmaster
4.1
4
4.2
4.3
4.7
4.2
4.3
3.9
4.2
4.3
4.3
4.2



## IBM Datacap Taskmaster Capture 8.01

- Advanced capture capabilities for the network or the browser
- Production Document Capture
  - Central (corporate or mailroom-type)
  - Departmental
  - Web-based distributed deployments
- Open Rules Processing System empowers the end user to orchestrate capture systems with a point and click interface
  - Input from scanners, MFDs, folders, Fax, email and electronic
  - Image enhancement to optimize recognition
  - Document ID and classification
  - Data extraction using OCR, ICR, OMR, barcode 1D and 2D technologies
  - Validates data including lookups (and retrievals)
  - Intuitive User Interface for data entry, edit and verification
    - Learns from human operator handling of exceptions
    - Thin and thick clients have same functionality
- Rich administration tools
  - Web-based monitoring and administration
  - Extensive Reporting

# Support for IBM ECM Repositories Standard

- IBM FileNet Content Manager (P8)
  - Tight handoff of data and images
- IBM Content Manager (CM8)
  - Uploads PDFs or TIFFs
- IBM FileNet Image Services
  - Uses IBM FileNet IDM Desktop Client
- IBM Production Imaging Edition
  - One product for the complete lifecycle of imaging from capture to repository to BPM to sophisticated viewer
  
  - Plus...
- IBM FileNet Capture Professional
  - Add automation and thin client distributed capture to existing IBM FileNet Capture installations

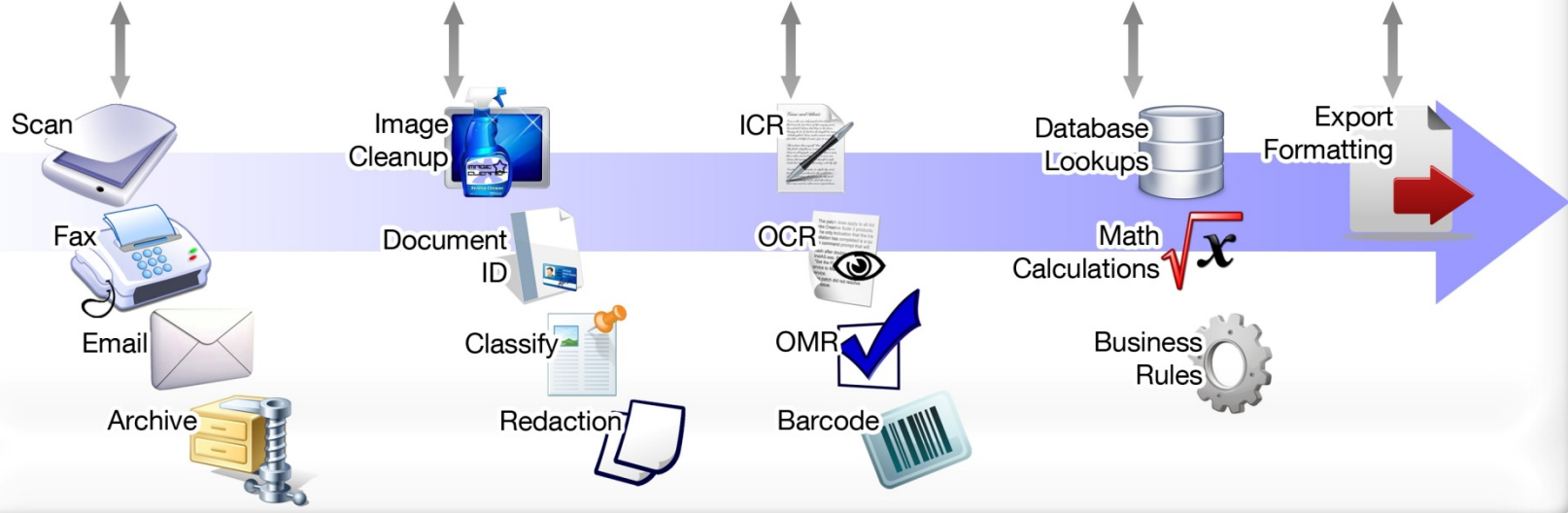
# IBM Datacap Taskmaster Addon Components

- Datacap Taskmaster Accounts Payable Capture
- Datacap Taskmaster Medical Claims Capture
- IBM Datacap Rulerunner Enterprise - for increased performance
  - Support for multi processor systems
  - Server-based Fingerprint processing
- Connectors for eMail and Electronic Documents
  - Provides ability to convert Microsoft Word, Excel, Outlook, PDF, and multipage TIFF files to single page TIFF files for capture processing
- Connector for Fax
- Connectors: EMC Documentum, Microsoft SharePoint and OpenText LiveLink
- Recognition support for major European languages (single-byte, Latin 1-4 code pages)

- Accounts Payable
- Claims Processing
- Healthcare
- Insurance
- Government
- Financial Services
- Transportation
- Manufacturing

- Browser-Based Scan Panel
- Thick Client Scan Panel
- Browser-Based Verify Screen
- Thick Client Verify Screen
- Administrator Job Monitor

# Rulerunner Service



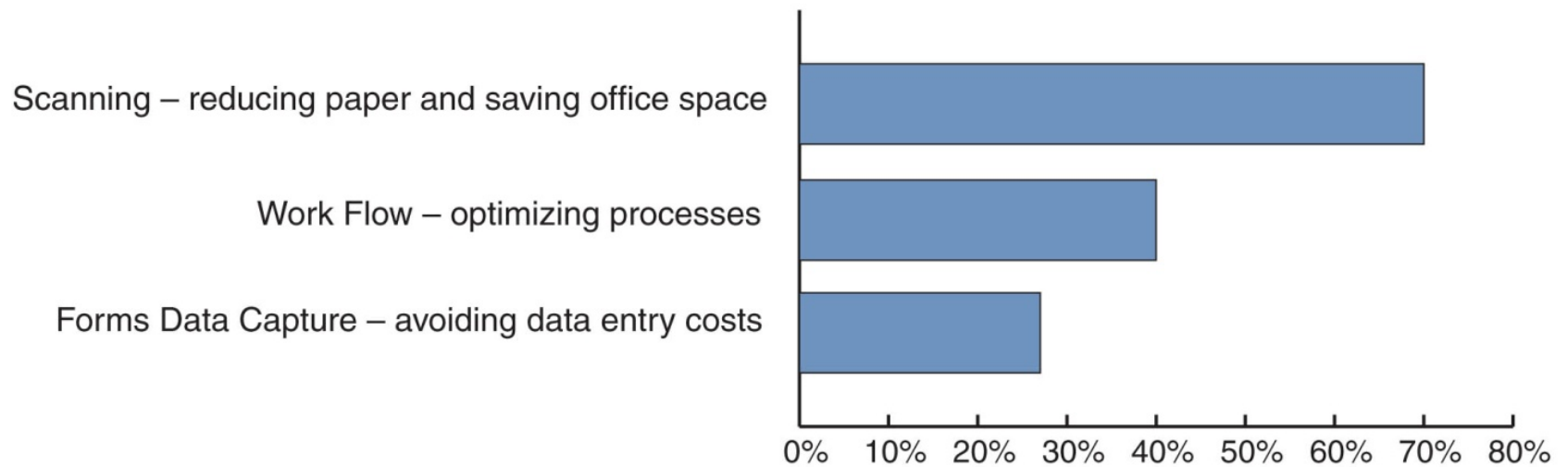
The screenshot shows the Datacap Studio interface with three callout boxes:

- Action in a Rule:** Points to the 'LookUp' action within the 'Rule: TaxpayerName' in the Rulesets pane.
- Action Library:** Points to the 'LookUp' action in the Actions library pane.
- Rule Bound to Field:** Points to the 'LookUp : Rule: TaxpayerName' entry in the Document hierarchy pane.

The interface includes a Document hierarchy pane on the left, a Rulesets pane in the center, and an Actions library pane on the right. The Rulesets pane shows a rule named 'Rule: TaxpayerName' with a function containing 'OpenConnection', 'ExecuteSQL', and 'PopulateWithResult' actions. The Actions library pane lists various actions such as 'equalize', 'Export', 'LookUp', 'CloseConnection', 'ExecuteSQL', etc.

# Where Is the ROI?

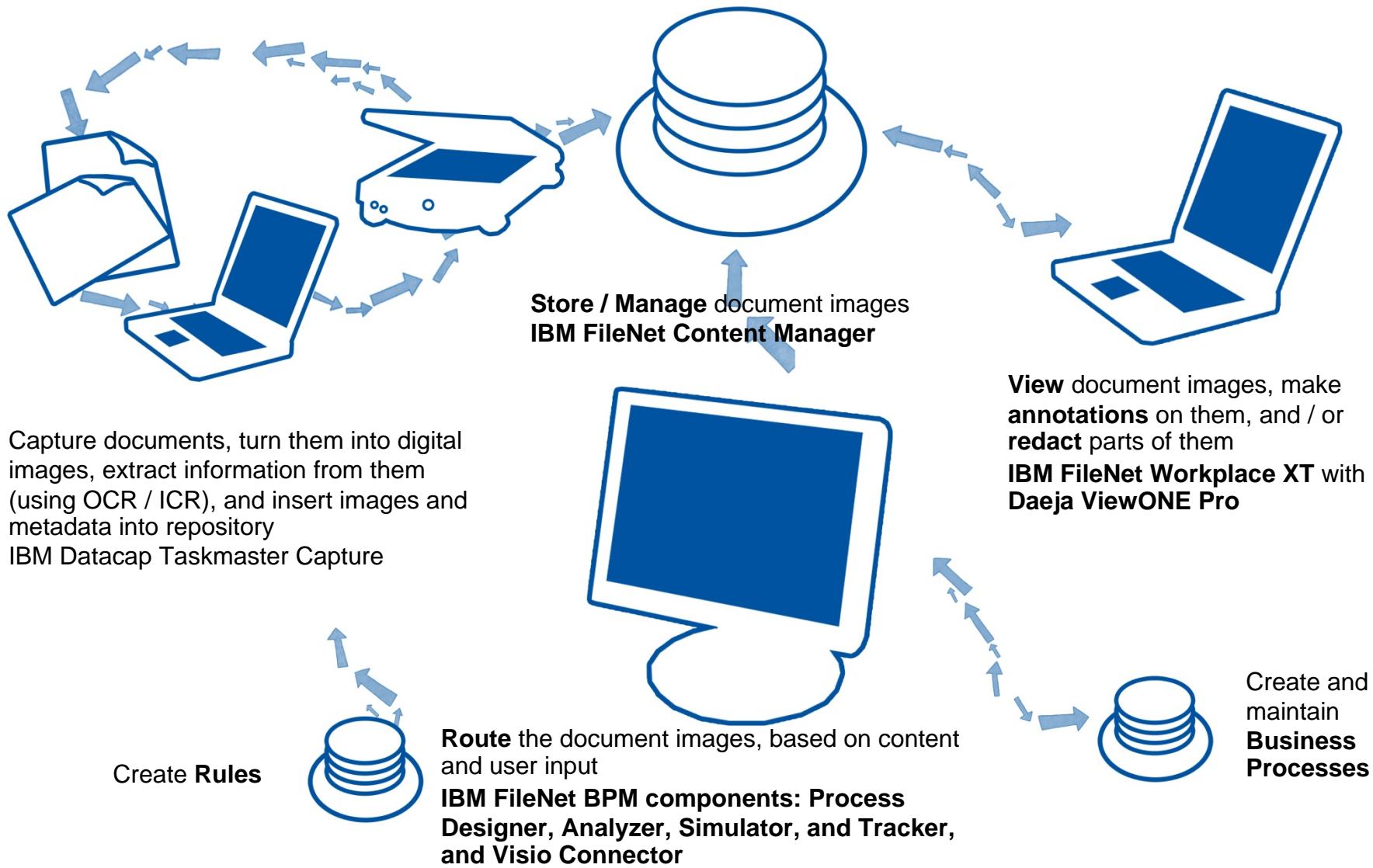
*Which aspects of scan and capture have produced the highest return on investment?*



Capture and Business Process. AIIM Market Intelligence. 2010



# IBM Production Imaging Edition



# Q & A

Learn more

IBM Datacap Taskmaster Capture

<http://www-01.ibm.com/software/ecm/datacap/taskmaster/>

IBM Production Imaging Edition

<http://www-01.ibm.com/software/data/content-management/production-imaging/>

Craine Communications Group

<http://crainegroup.com/>