Enterprise Content Management 2012

UserNet Forum - Istanbul

Datacap Taskmaster Capture Update

David Jenness, IBM ECM, Document Imaging and Capture

Master The Big Picture

- Smarter Document Capture with **Datacap Taskmaster**
- Active Document Imaging with **Production Imaging Edition**

The Document Automation Evolution

The Document Touchpoints:

- Arrival and Batch Creation
- Document ID and Sorting
- Data Entry
- Validations
- Routing

Today, every one of these steps can be automated.

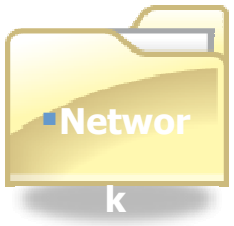
It starts with Multi-Channel Input

- High speed production scanner

- Desktop scanner

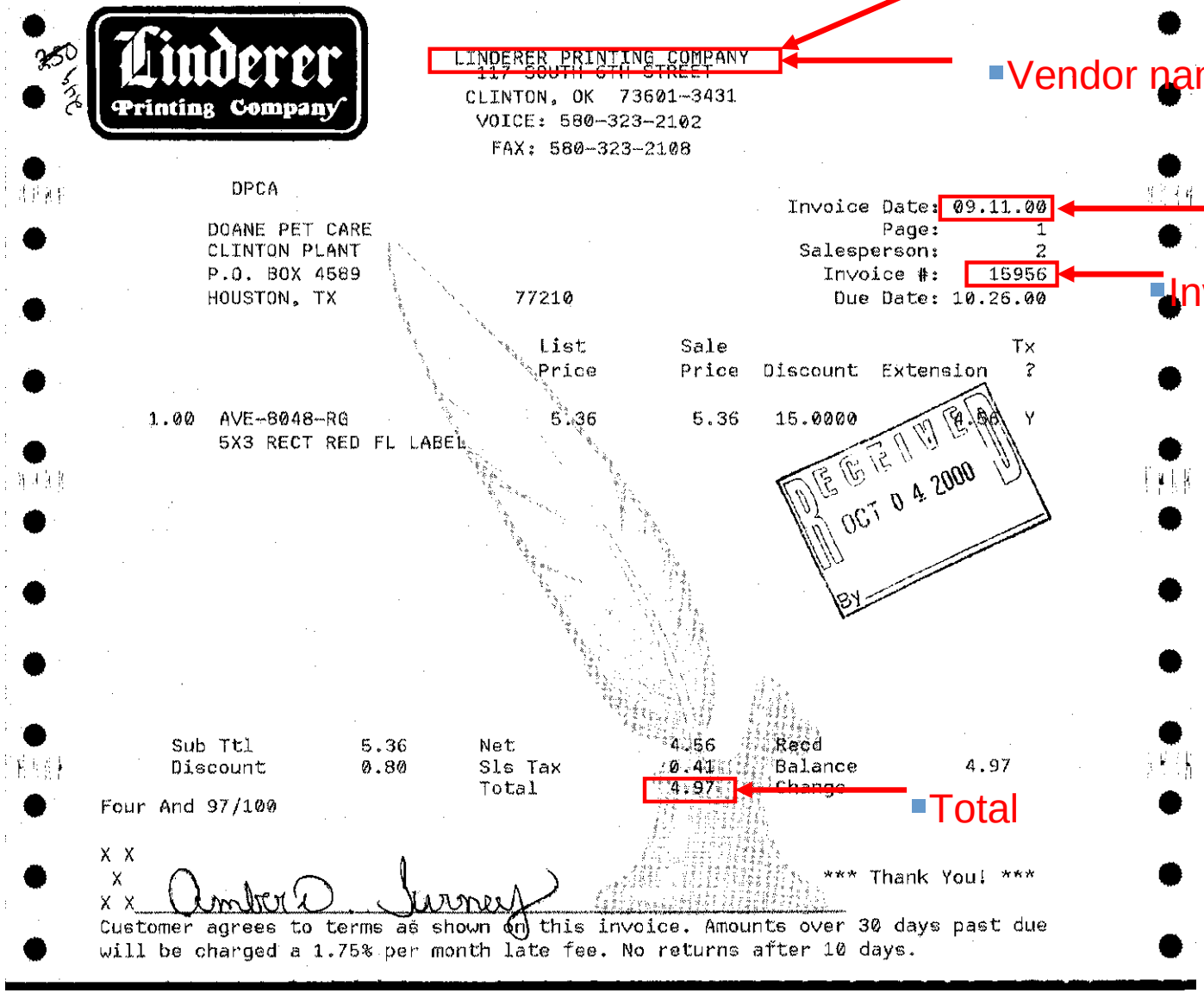
- Multi Function Peripheral (MFP)

- Fax



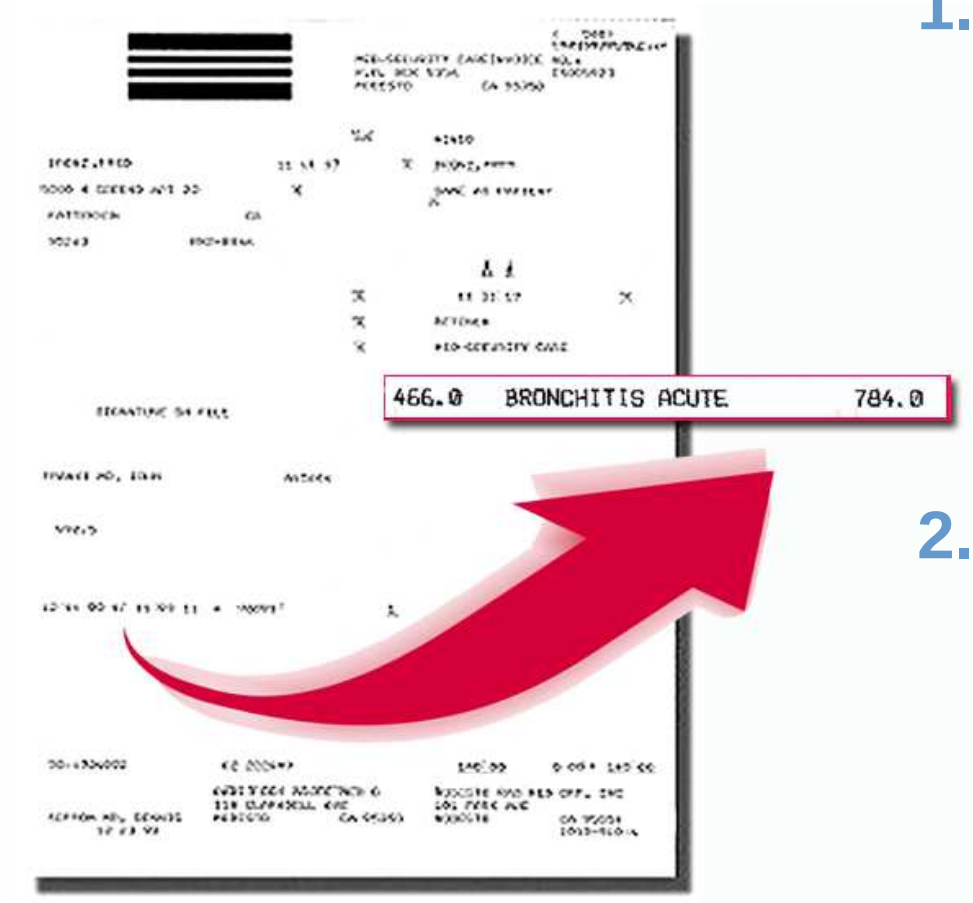
- Emails and attachments

Optical Character Recognition (OCR)



- Vendor ID lookup
- Vendor name
- Invoice date
- Invoice number
- Total

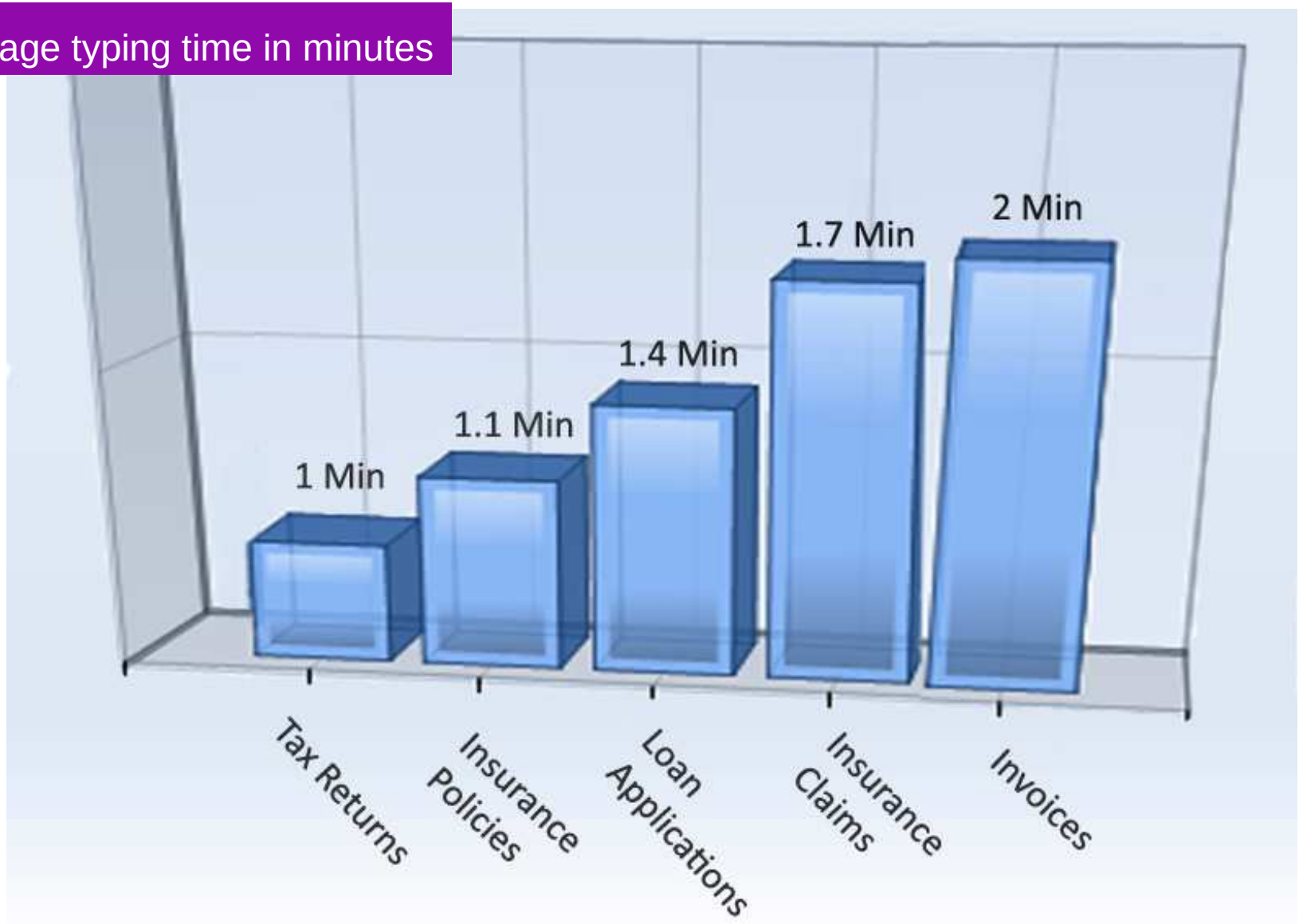
Why Extract Data?



1. **Index the document image with searchable metadata for ECM**
   - Vendor name
   - invoice number
   - invoice date
   - SSN
   - Customer number
   - Provider ID
2. **Populate Business Systems with Key Data**
   - ERP
   - Accounting system
   - CRM
   - Workflow
   - Adjudication
   - Order processing
   - Business Intelligence

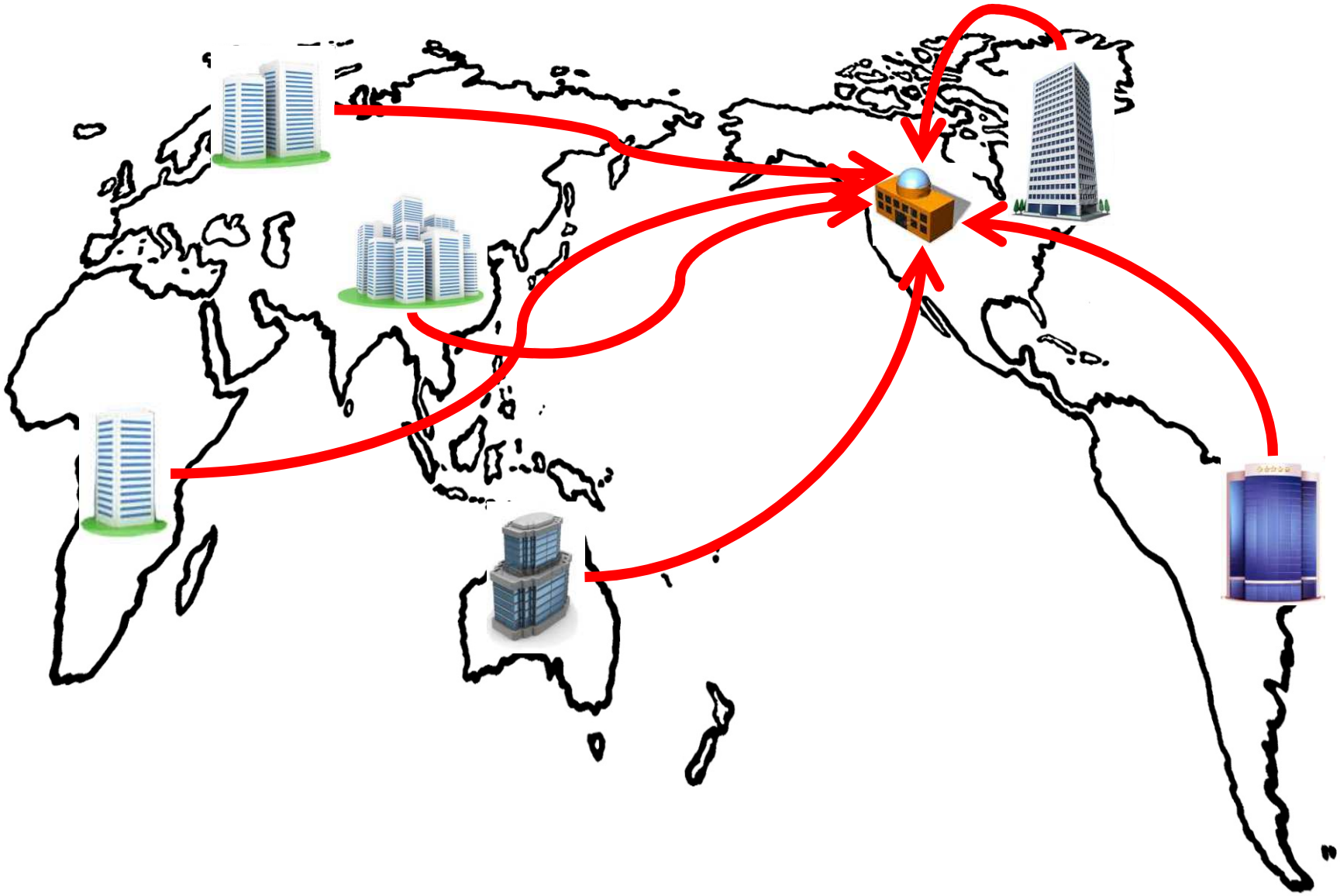
Data Entry Time Spent Per Document

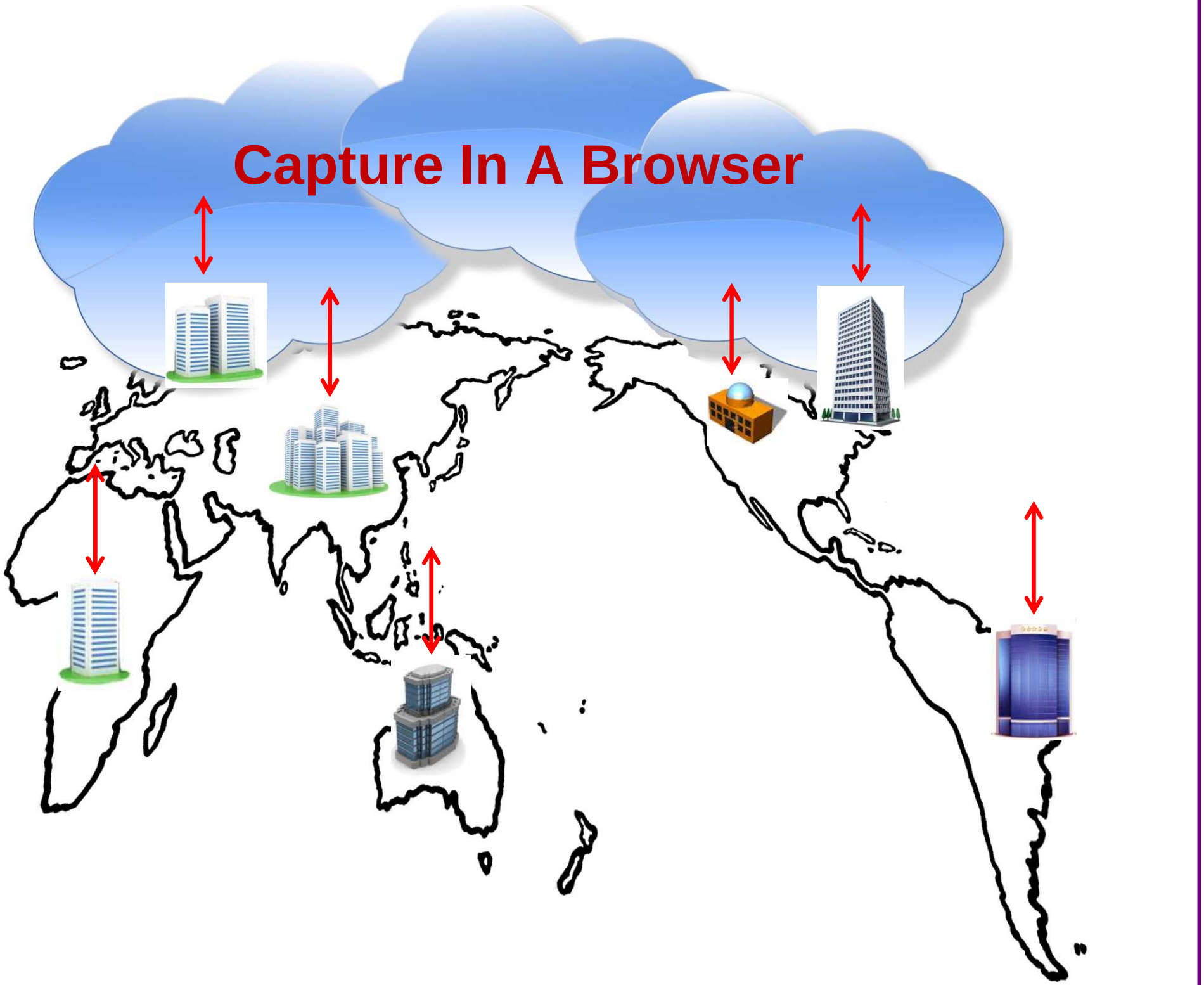
- Average typing time in minutes



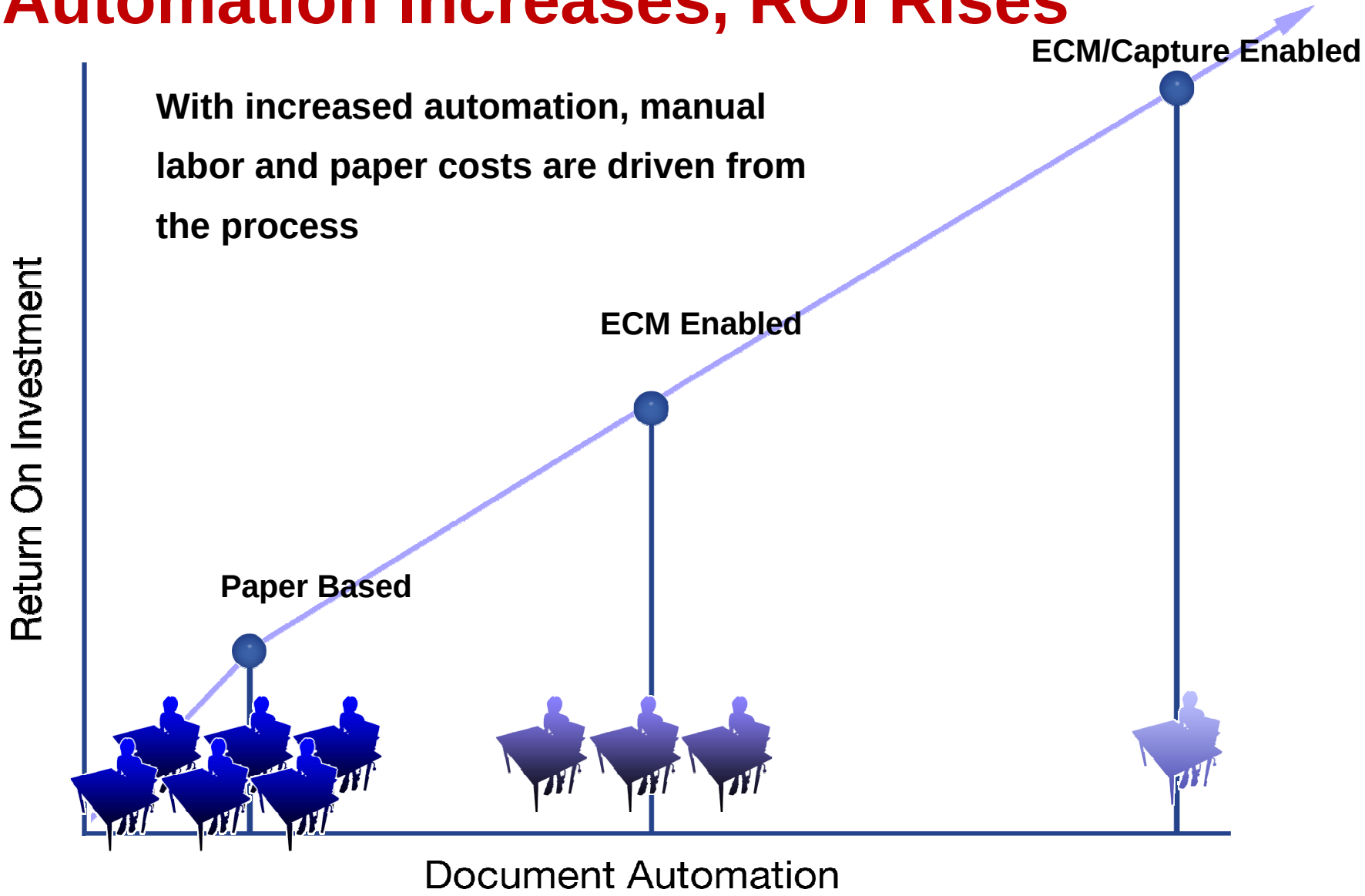
- *Sources: Health Data Management, IOMA, Datacap*

Paper Documents Shipped To Processing Center




Capture In A Browser

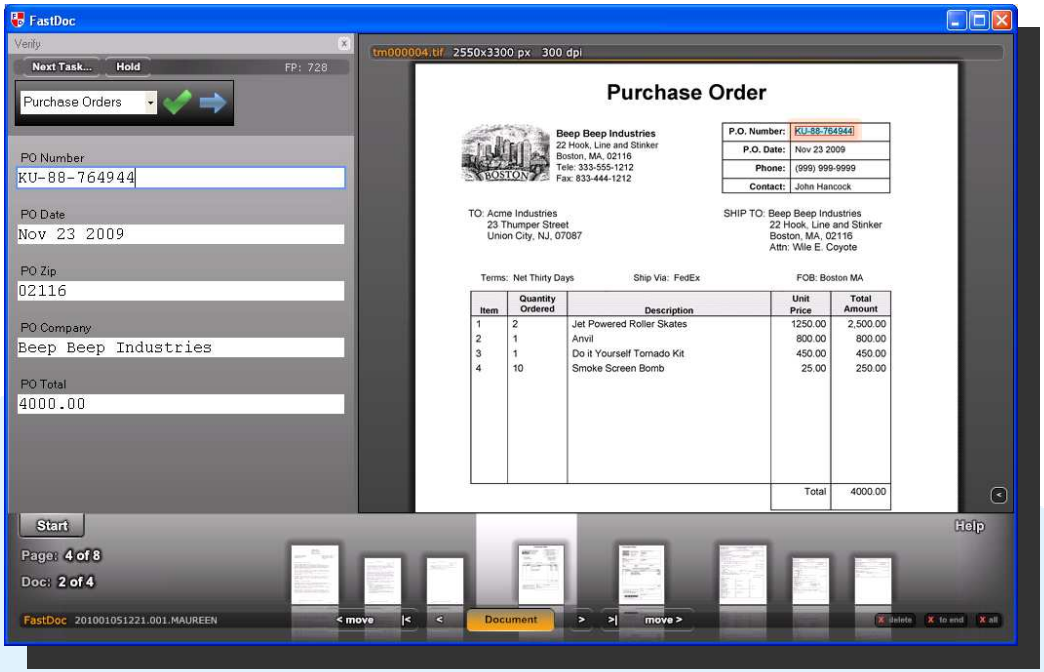
As Automation Increases, ROI Rises



Business Benefits of Intelligent Document Capture

- Accelerate document input
  - Deliver products and services to customers faster
- Eliminate manual data entry
  - Reduce headcount doing low value tasks
  - Eliminate typos and lost documents
  - Other cost savings
  - Eliminate document shipping
  - Eliminate paper handling
  - Eliminate paper storage
  - Control who has access to documents
  - Maintain compliance

Datacap Taskmaster Capture



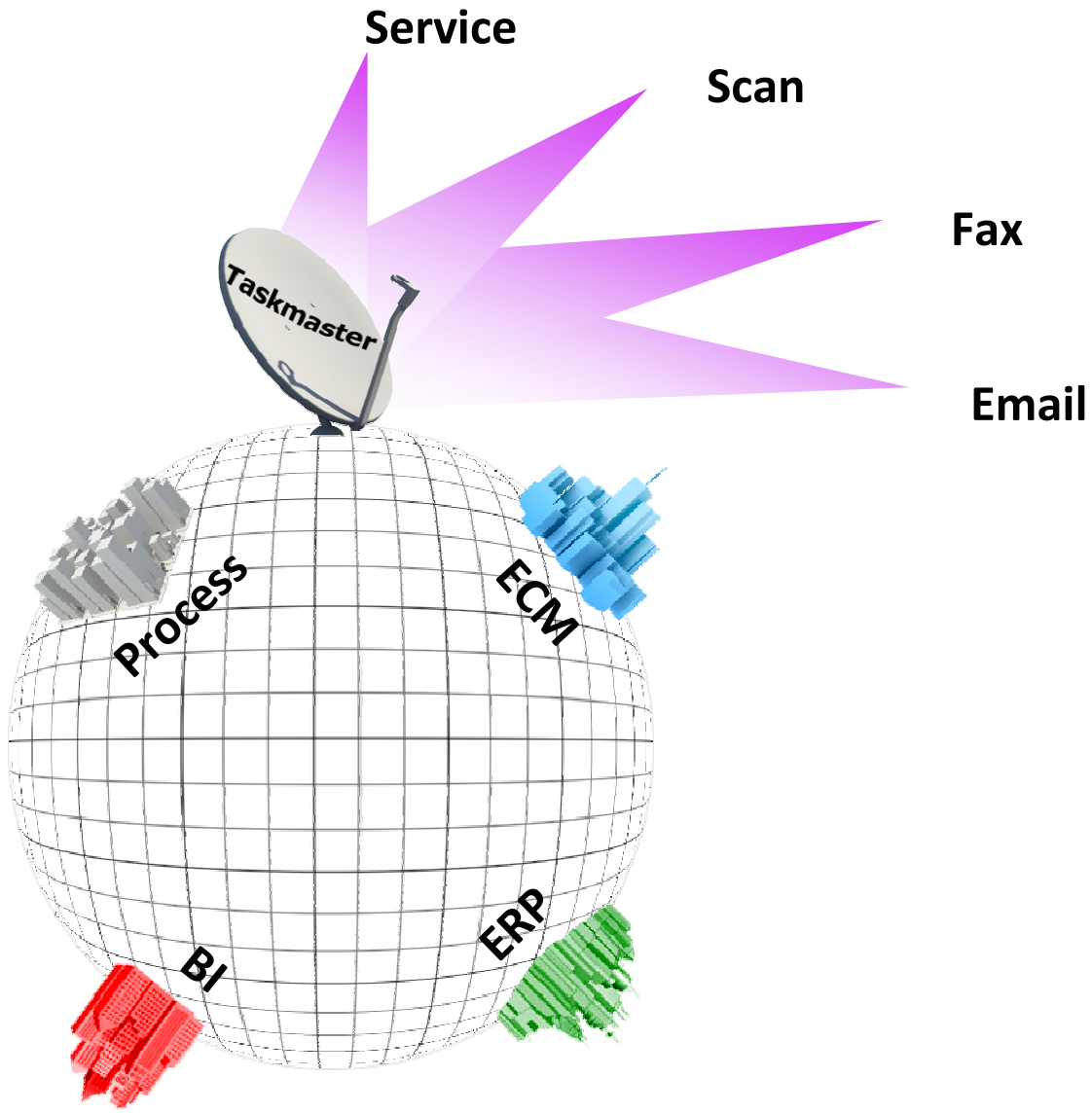
- **User Interface presents single, attractive panel for Scan and Verify**
- **Highly customizable, rules-based environment – 3 leading OCR engines, plus Parascript, Content Classification, 100s of validations, export to all IBM repositories plus 3rd parties**

- Three Pricing Models
  - Taskmaster Entry Level
  - Taskmaster Capture
  - Taskmaster Enterprise Edition
- Global language support (Pan-European, Cyrillic , Simplified Chinese)
- Rapid Application Deployment environment
- SAP integration with XFT
- User-based licensing

The Big Picture for Capture

Multichannel capture for the Enterprise

- Documents arrive through many conduits
- Set a single standard for all incoming documents
- A capture portal
- One solution to convert, “understand,” and tag all incoming documents
- Then direct them to the right systems, tools and people



Datacap Taskmaster (Entry Level) → **Datacap Taskmaster Capture** → **Datacap Taskmaster (Enterprise Edition)**

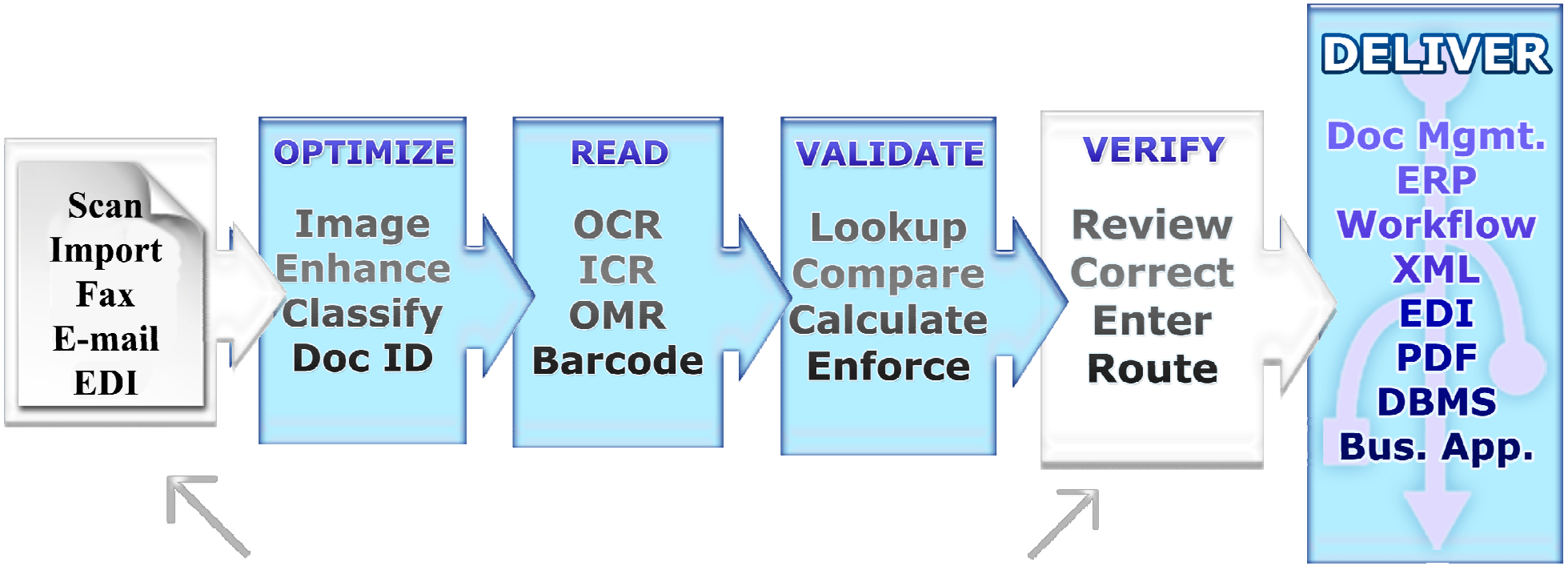
Backfile conversion to reclaim storage space	Real time scanning anywhere anytime	Accounts Payable capture for supply chain	Tax form processing for govt. revenue	Mortgage capture to accelerate lending
Basic scan and imaging	Capture for SAP - HR and contract management	Medical claims capture for timely payments	Survey/Census processing for accurate data gathering	Capture for customs clearance
Capture for SharePoint* *with ECM Connector	Medical Records capture for better care	Capture for Subscription processing	New customer onboarding for fast service	Enterprise capture
Scan and archive for IBM Content Manager	Capture fax, email and attachment	Correspondence capture and tracking for compliance	Sales order entry for fast fulfillment	Custom applications

Datacap Capture Community Perception

	AnyDoc Software OCR for AnyDoc	Autonomy Cardiff Teleform	Brainware Distiller	IBM FileNet ADR	IBM DataCap Taskmaster	Kofax Transformation Modules	ReadSoft FORMS	ReadSoft INVOICES	Top Image Systems eFLOW
Ease of Use					4.1				
Ease of Roll Out/Scalability					4				
Ease of Configuration/ Customization					4.2				
Security					4.3				
Strategy					4.7				
Stability/Bugs					4.2				
Integration					4.3				
Community and Channel					3.9				
Support/Maintenance					4.2				
Value for Money					4.3				
Standards					4.3				
Application Development					4.2				

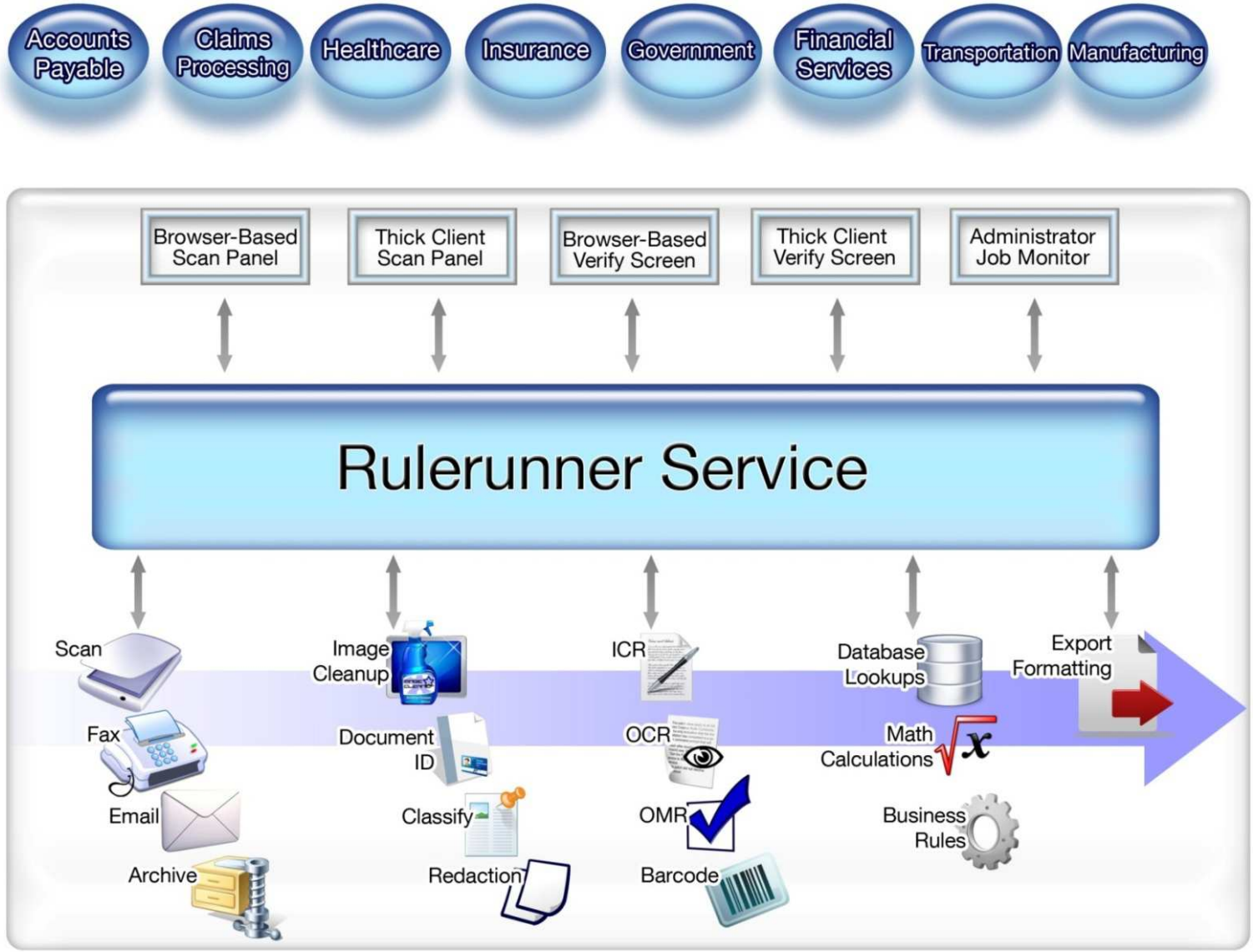
- In the 2011 AIIM Capture Software Product Study, no product scored higher than IBM Datacap Taskmaster Capture – well above Kofax and EMC Captiva

The Capture Process

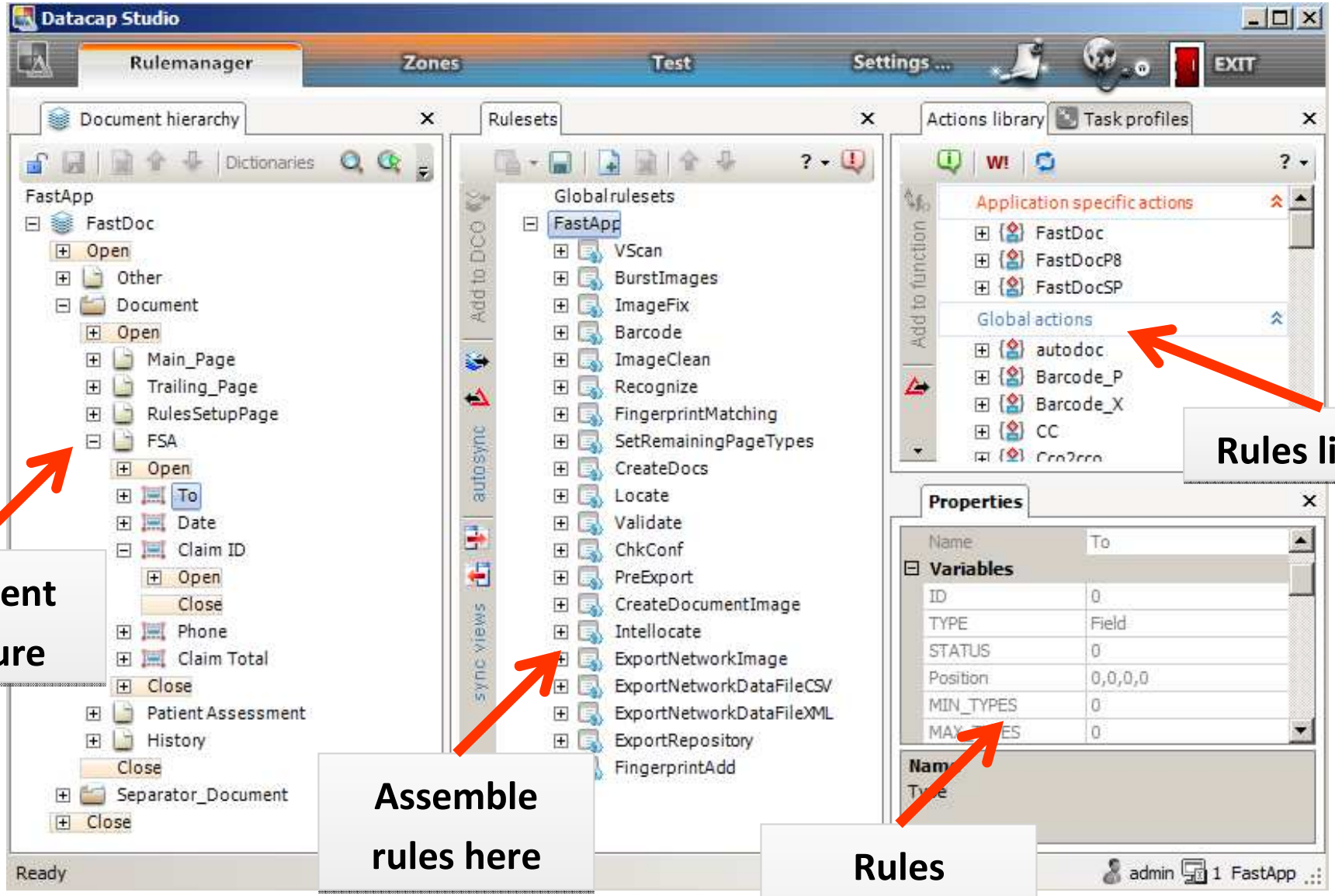


Browser Based Capture

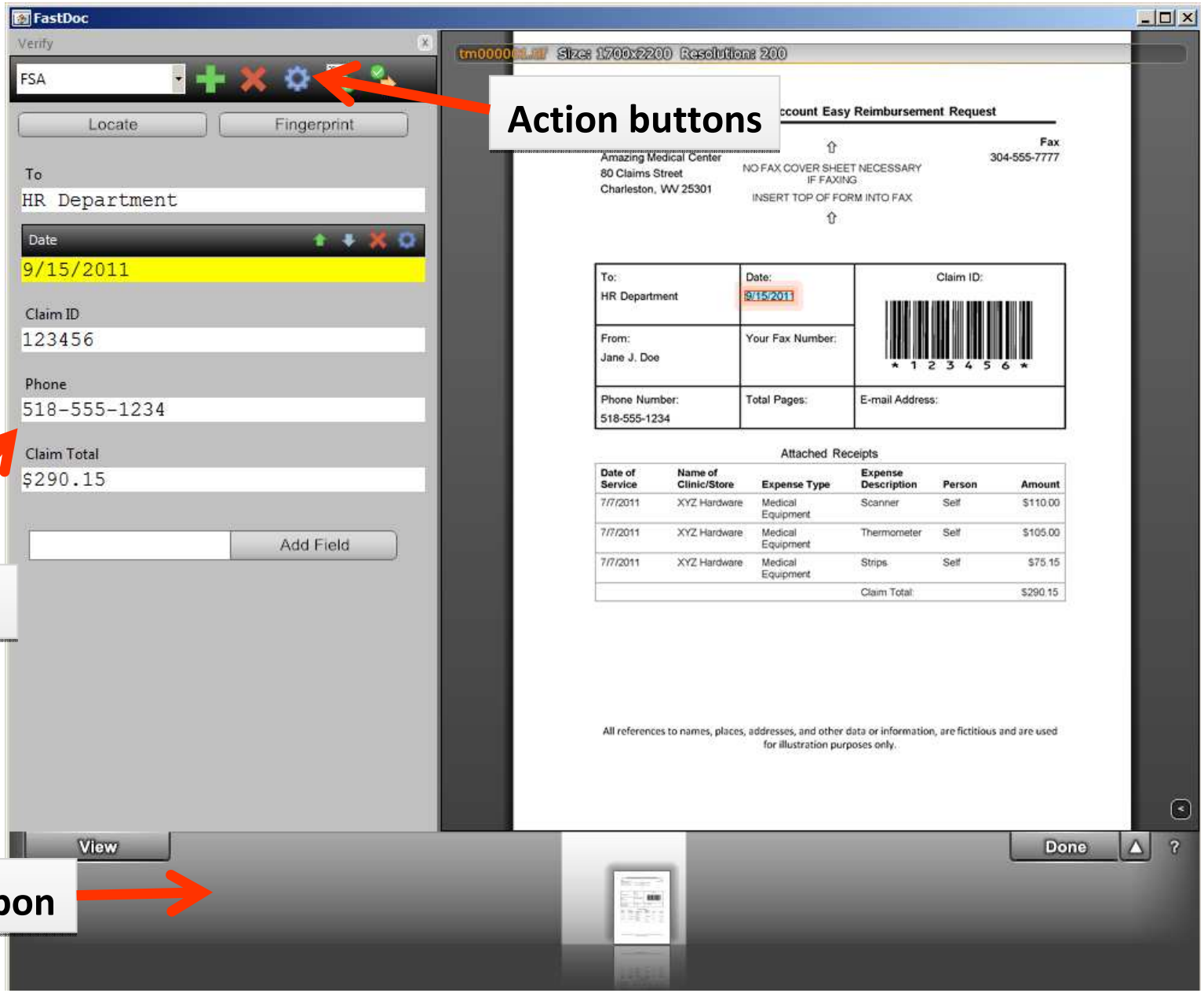
Stability: Robust, Flexible, Scalable



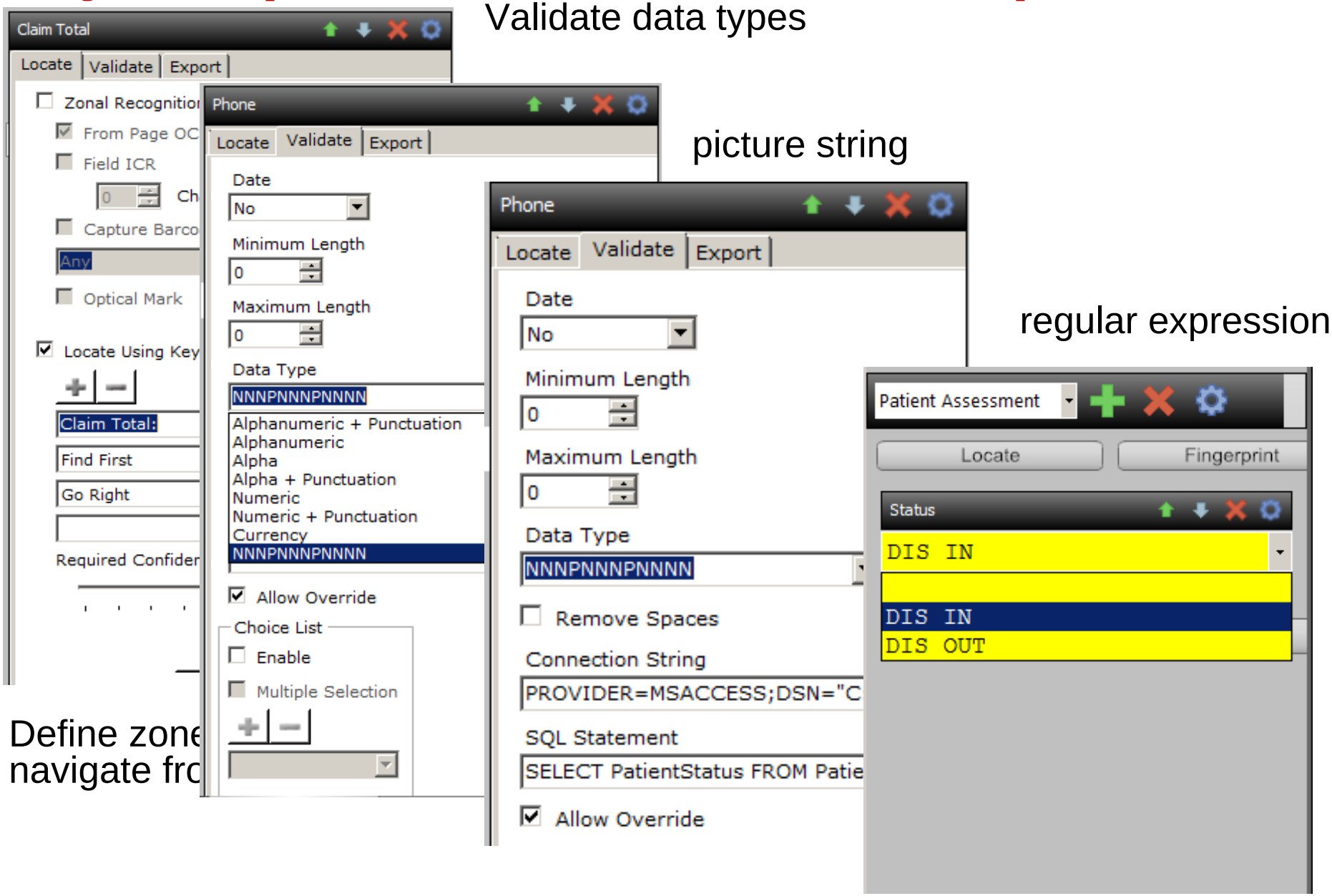
Datacap Studio for configuration



Simple, easy-to-use user interface



Easy setup for Automation Techniques



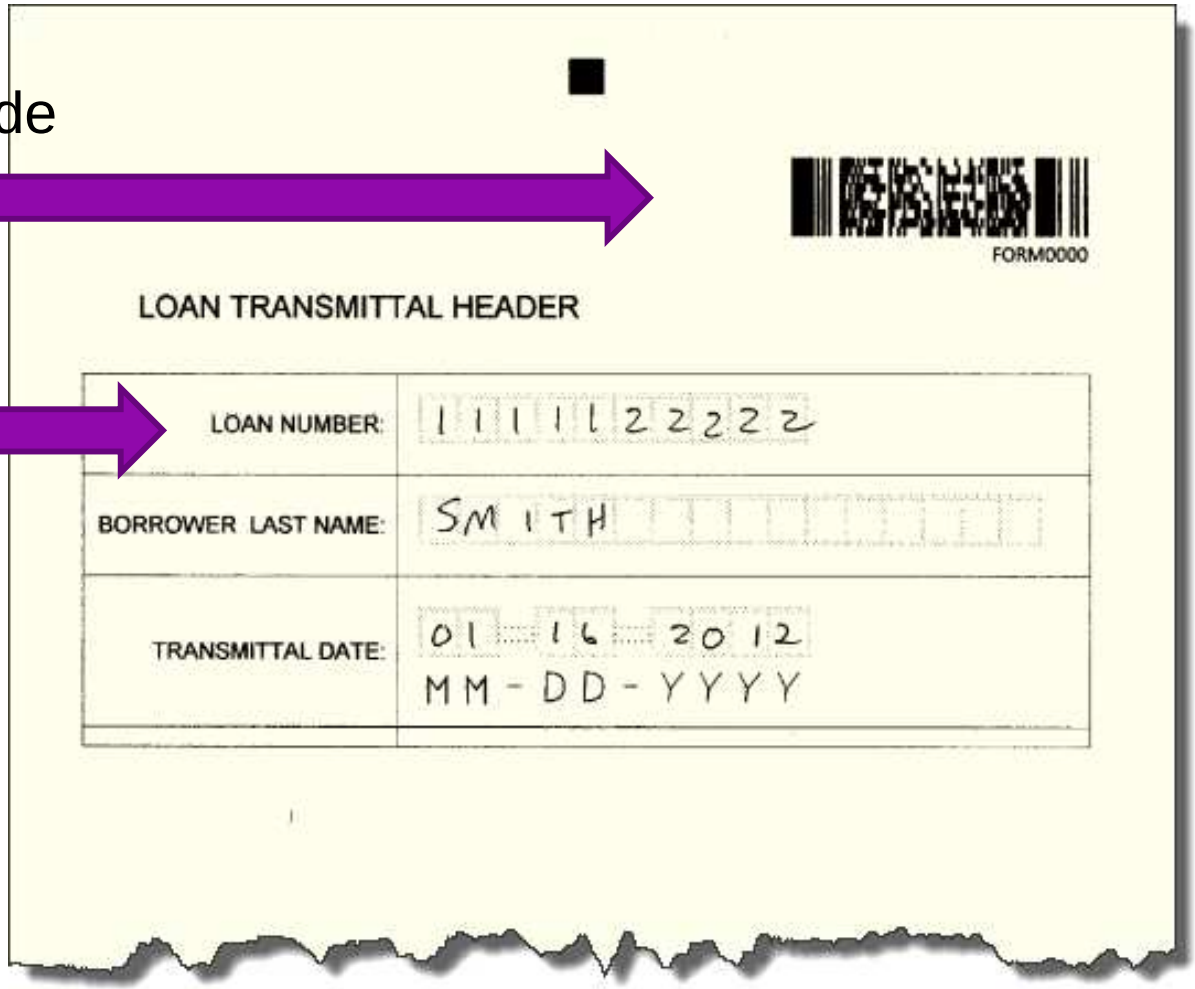
The Last Frontier…

Document Classification.

The identifying and sorting of document types to create batches….

Smart Separator Sheet

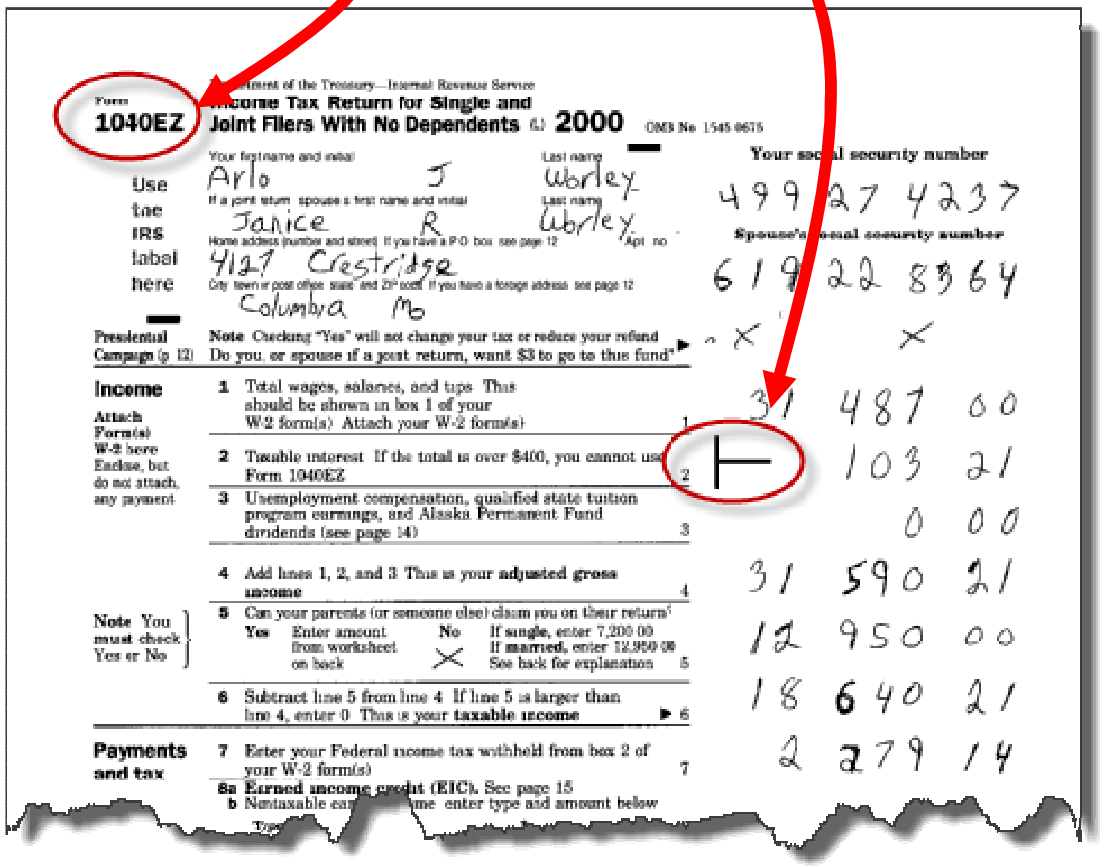
- Document / Form type barcode
- Additional Data
  - Could be pre-printed
  - Or entered by user



- **Shortcomings:**
- Labor-intensive
- Rely on a person's knowledge to correctly identify the documents

Pattern Recognition

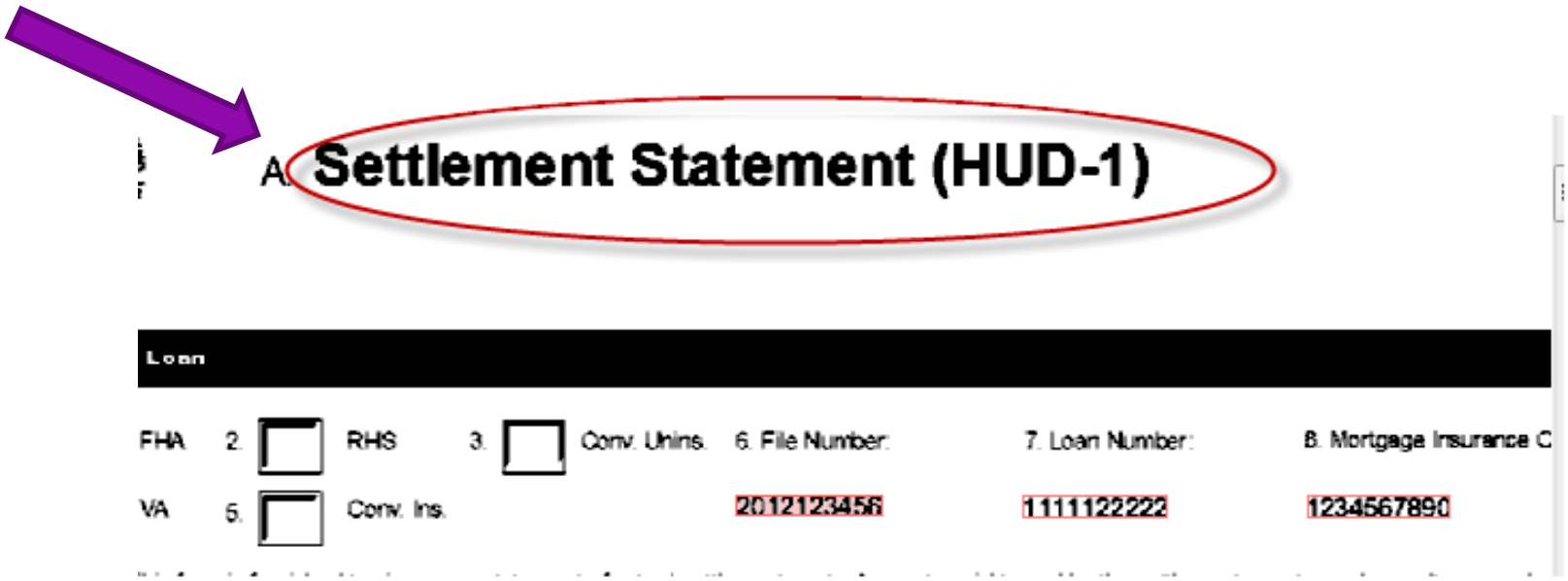
- Very fast matching to unique marks on a page – “anchors”
- Used with fixed forms
- Most commonly used with ICR – handprint forms
- PatternMatchIdentify Action



Keyword

- Following OCR to recognize machine print text on a page
- Regular expressions find key words and phrases
- Search zones or search the entire page
- Searches can be stored externally in key files
- FindRegexList or FindRegexList_InZone

\bSettlement\s*Statement.*HUD.*[1]\b



Fingerprint Recognition

- Fast (sub-second) – does not require OCR
- Matches the patterns of light and dark - Characters, blobs, words, text lines
- Supports thousands of stored page templates
- Also differentiates between multiple formats of the same page type
- Adjusts the positions of zoned fields

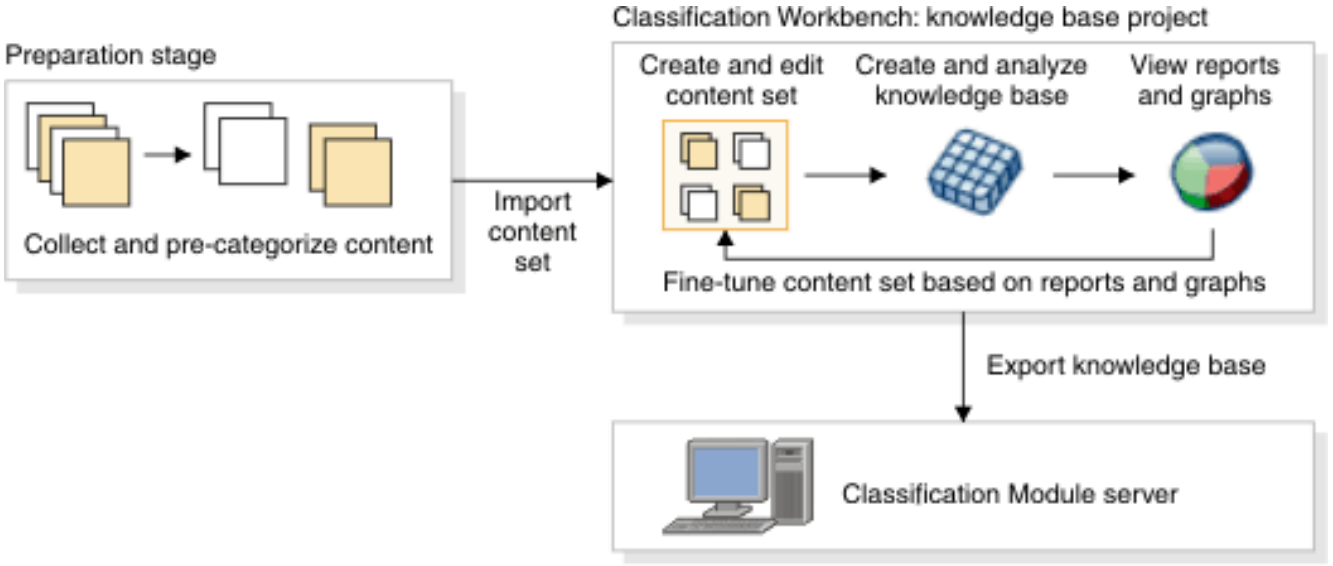
Scanned Image

Comparing patterns of light and dark

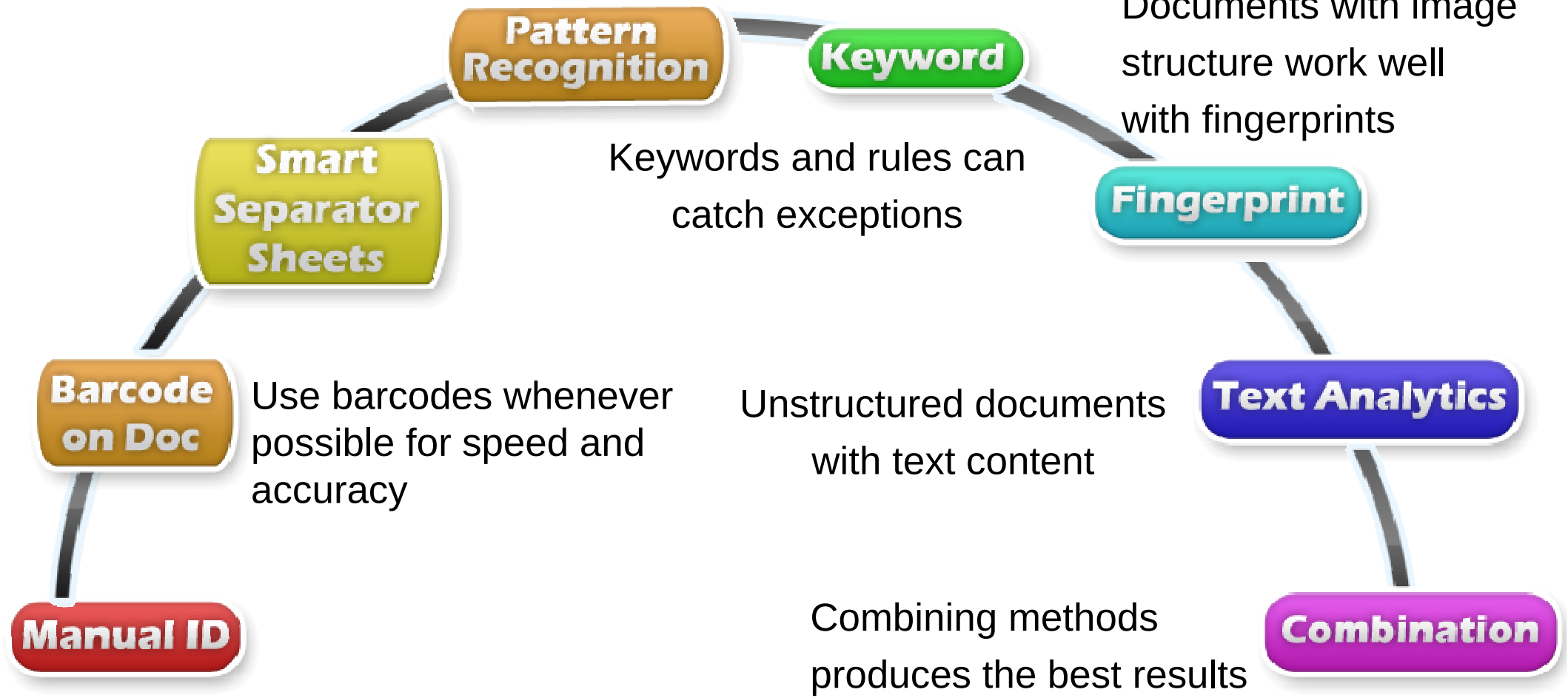
Fingerprint

Text Analytics Concepts

1. Collect and organize sample content
2. Create, analyze, and learn
3. Assess performance, review reports
4. Publish to the Content Classification server



Taskmaster Capture Classification Spectrum



Why Smarter Capture?

- **Convert:** Turn paper, fax, email, electronic attachments into images
- **Classify:** Understand the nature of the document and its role
- **Extract:** handprint, typed text, handwriting, bar code, check box
- **Assure Quality Data:** database lookups, math calculations, P.O. matching
- **Protect:** Redact fields, split and rendezvous for privacy
- **Format:** Both image and data for ECM, ERP, web, and business systems
- **And Do it in *Your Language*!**

Understand Your Documents

- Document type
- Data field location
- Retention schedule
- SSN, Customer/Vendor ID
- Date received
- Document number
- Business rules
- Exception Handling

Manual Approach	Smarter Approach
Print out and scan faxes, emails and attachments	Convert fax and electronic documents to image
Manual data entry	Automatically extract data from image
Manually compare name to existing table	Automatic database lookup and match
Manual Doc ID and sorting	Automatic document ID with analytics

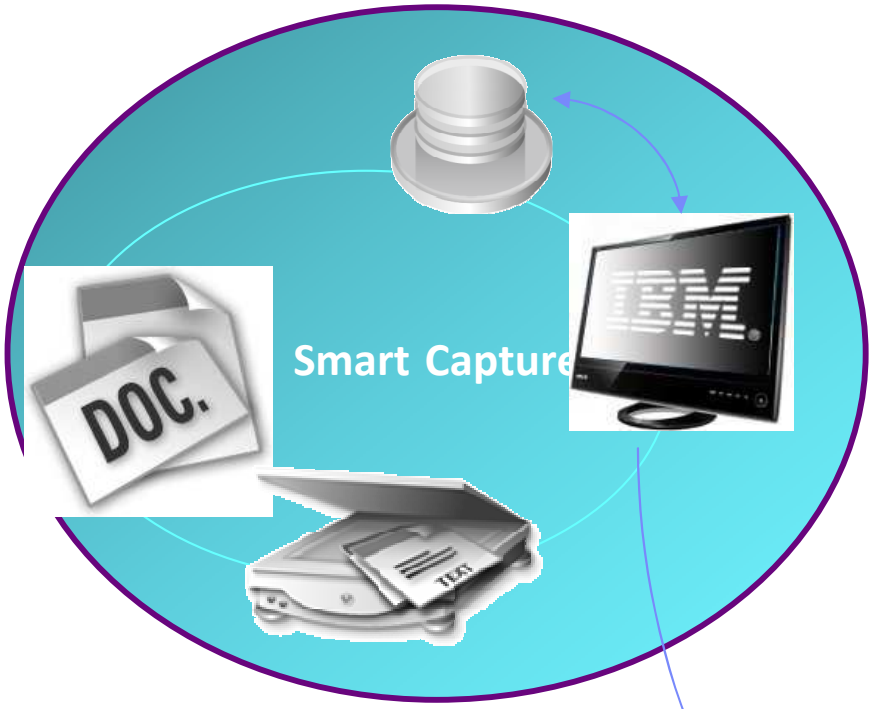
What is IBM Production Imaging?

PIE bundles together state-of-the-industry imaging solutions to enable sophisticated document process automation.

- **Input – multi-channel, central and distributed**
- **Storage, retrieval, compliance, disposal**
- **Workflow –view, process, collaborate**

- PIE is a single product for scanning, OCR/ICR, document storage, annotation, redaction, viewing, and business process automation.
- Rapid ROI - capture, view, share, process documents anywhere, any time with less effort

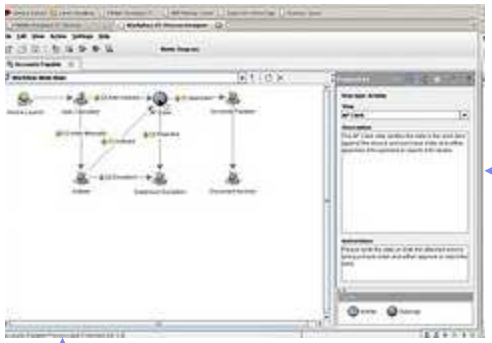
Production Imaging Edition



Input documents

Datacap Taskmaster Capture

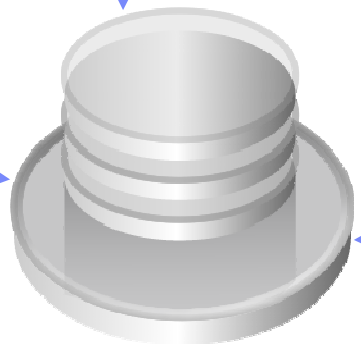
- "Understand" documents
- Aggregate documents
- Extract data
- Multi-Language
- Multiple formats



Route document images - based on content and user input

Imaging-centric workflow

with FileNet BPM



Store / Manage document images

IBM FileNet Content Manager

View document images - make *annotations* on them, and / or *redact* parts of them

IBM FileNet Workplace XT

with **Daeja ViewONE Pro**

IBM Production Imaging Edition

Automate the life cycle of a document, from capture to business process to archive

The Total Intelligent Imaging Environment

- Smarter Capture – central and remote
- Flexible workflow based on document content
- Access by permission
- Set retention schedules
- Annotate and redact documents

Configurable:

- Set up capture applications with a point and click interface
- Build document-centric workflows with visual mapping
- Adapt to your Business Rules

Benefits of Intelligent Imaging:

- Eliminate paper
- Extract data from paper
- Route data to Line of Business systems
- Control distribution of documents
- Enable communication about documents
- Reduce cost

Enterprise Content Management 2012
UserNet Forum - Istanbul

The Seven Secrets of Success for Capture Engagement

An Exclusive Luncheon and Demonstration Event at IOD for Capture and PIE users

Mandalay Bay, Las Vegas
Monday, October 22, 2012
12:30 - 2 PM
ECM Central

Sponsored by FUJITSU

Limited seating!

RSVP: deeptnag@in.ibm.com

Enterprise Content Management 2012

UserNet Forum - Istanbul

Thank You

Presenter name here