


# Avoid Paper cuts WITH SHAREPOINT

## Turning paper into gold

Chris Riley, ECMp, IOAp, v-TSP  
Sr. ECM and Document Imaging Architect  
Twitter: @rileybeebz  
ShareSquared, Inc.

# Why Chris?

- I know ECM and SharePoint
  - ECM Author
  - Certified  ecmP, ioaP  
aiim practitioner, aiim practitioner
  - Past President AIIM GG, AIIM National Board
  - “Imaging savant” - leading expert in document imaging
  - Microsoft v-TSP, Technical Solutions Professional
  - Sr. ECM & Document Capture Architect at full-service SharePoint Consulting firm ShareSquared, Inc.

# ShareSquared, Inc.

- SharePoint MVPs
- MS Certified Trainers
- Former SharePoint Dev. team
- Microsoft Managed Gold Partner
- Full-Service SharePoint consulting firm



## Offerings:

- GamePLAN – Planning Services
- MeltDown – Improve existing deployments
- Hourly Consulting
- Composer + Maestro architecture tool

## Agenda

- What is document imaging
- How does SharePoint help
- Document Imaging Best Practices
  
- Hopefully not death by PowerPoint
- Not technical
- Free form session – keep it lively



# The Purpose

- Empower You
- Educate
- Give Best Practices
- To avoid the statistic
  - 70% of Document Imaging projects Fail!

# Dealing with paper

- Two approaches:
  - Converting physical to digital i.e. document imaging
  - Digital tracking of physical records
- Because:
  - Managing physical documents is HARD
  - Paper cuts are no fun
  - Unifying business processes would be cool
  - Paper is expensive and risky



# Dealing with paper

- Physical Records management
  - Driven by volume and storage costs
- Methodologies
  - Manual with Lists and workflows
  - Semi-Automated with Barcode and Label policies
  - Fully automated with RFID Technology
    - Third Party market only

# Barcodes and Lables

Automatic barcode  
information  
management policy

The screenshot shows a SharePoint site titled "Corprate Portal" with a "Contracts" list. A blue callout bubble points to a barcode on the document "Mutual\_NDA\_MS". The barcode is labeled with the number 8184875851. The document list includes columns for Name, Modified, Document ID, Rating (0-5), and Taxonomy.

Type	Name	Modified	Document ID	Rating (0-5)	Taxonomy
	Confidentiality Agreement	10/6/2010 2:12 PM	CORDOCS-7-1	☆☆☆☆☆	Contracts:Contract Template
	Contract between Ted and Chris	11/9/2010 1:13 PM	CORDOCS-7-2	☆☆☆☆☆	Contracts:Executed Contract
	Cooperative_Agreement	11/9/2010 1:30 PM	CORDOCS-7-3	☆☆☆☆☆	Grants:Grant Submission
	Mutual_NDA_MS	2/20/2011 11:33 AM	CORDOCS-7-20	☆☆☆☆☆	Contracts:Non-Disclosure Agreement
	Policy Acknowledgement	10/8/2010 10:47 AM	CORDOCS-7-7	☆☆☆☆☆	Contracts:Executed Contract
	schoolagreement	2/20/2011 11:10 AM	CORDOCS-7-6	☆☆☆☆☆	Contracts:Non-Disclosure Agreement
	SPLG	12/13/2010 1:28 PM	CORDOCS-7-4	☆☆☆☆☆	Contracts:Executed Contract

BUT!

SharePoint Does not have document imaging OOTB

And that is a good thing!

# Document Imaging

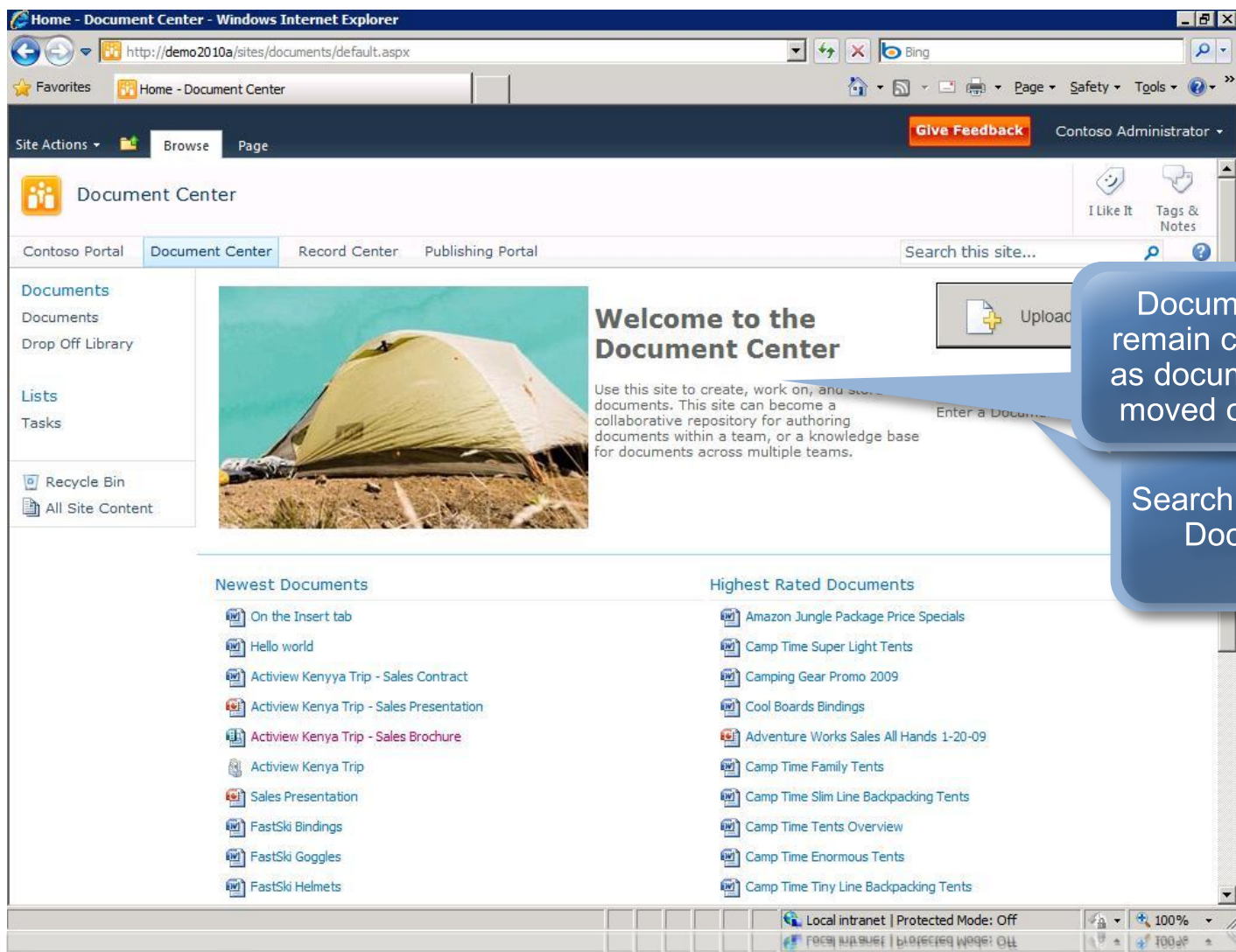
- As much about the setup of SharePoint as the imaging software you choose
- Reach for the stars, but don't expect magic
- Goal: Eliminate paper without data entry

## Key SharePoint Features

- Document ID Service
- Document Sets
- Content Organizer & Rules
- Remote Blob Storage (RBS)
- Conversion Services



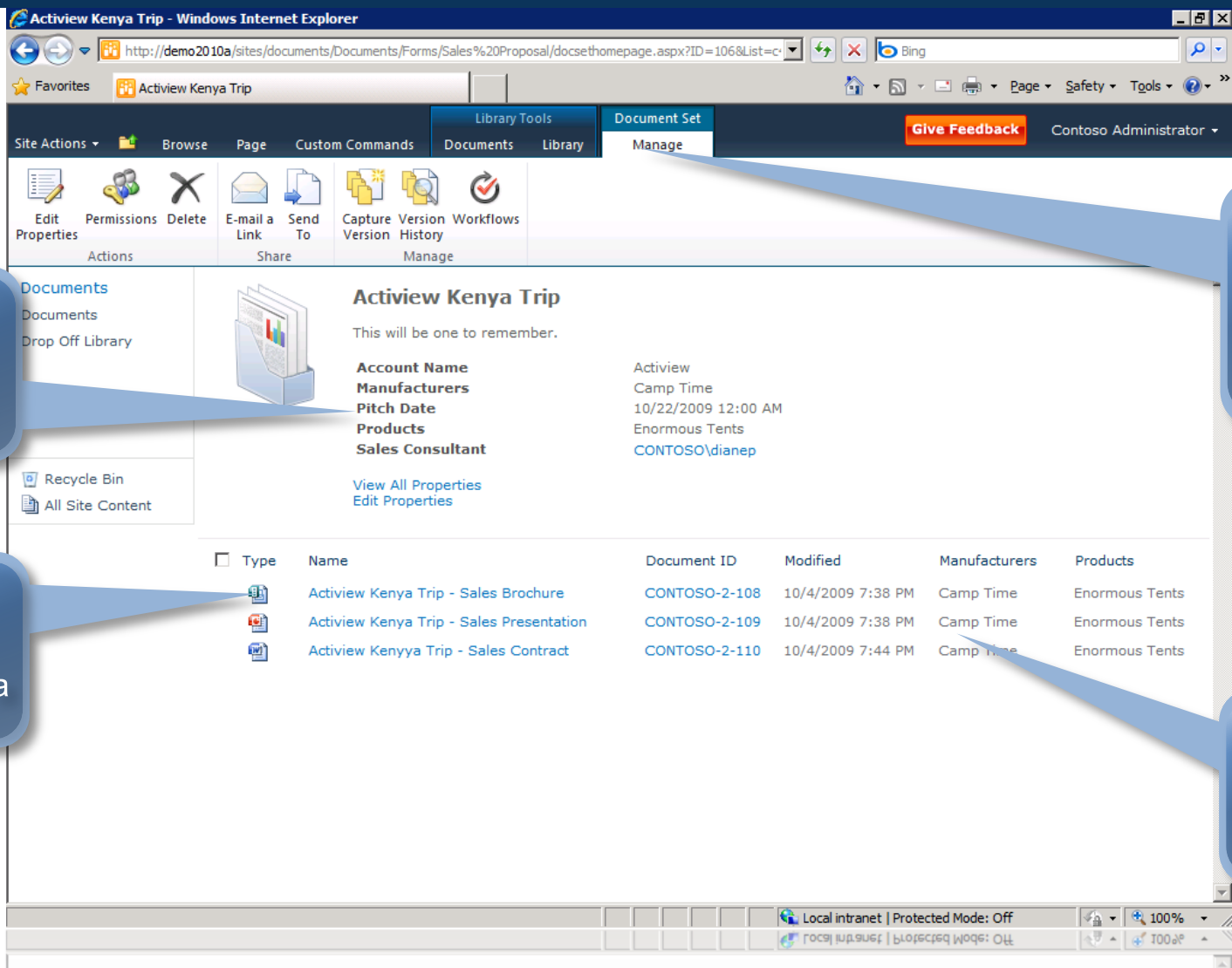
# Document ID Service



Document IDs remain consistent as documents are moved or copied

Search for a specific Document ID

# Document Sets



Document Set Properties

Manage Different Types of Content as a Unit

Document Set tab for Property, Version and Workflow Management

Individual Document Properties

# Content Organizer

Content Organizer Rules - Windows Internet Explorer

http://demo2010a/sites/documents/RoutingRules/Group%20by%20Content%20Type.aspx

### Content Organizer Rules: Edit Rule

**Submission's Content Type \***

By selecting a content type, you can specify the content type that will be assigned to the content when it is routed to a target location. In addition, submissions that match the conditions of this rule. In addition, submissions that match this rule will receive the content type selected here when they are placed in a target location.

Content type:

Group: Custom Content Types

Type: Sales Document

Alternate names:

This content type has alternate names in other sites:

Add alternate name: Sales Proposal Add

Note: Adding the type "\*" will allow documents of unknown content types to be organized by this rule.

List of alternate names: Sales Document Remove

**Conditions**

In order to match this rule, a submission's properties must match all the specified property conditions.

Property-based conditions:

Property: Confidentiality X

Operator: is equal to

Value: Restricted

(Add another condition)

**Target Location \***

Specify where to place content that matches this rule.

Destination: /sites/documents/Documents/Restricted

Example: /sites/DocumentCenter/Documents/

Automatically create a folder for each unique value of a property:

Select a property (must be a required, single value property):

Specify the format for the folder name: %1 - %2

When the folder is created:

Submissions matched to a content type

Property based conditions route content

Auto Provision folders based on properties or volume

Specify alternate content type names to accommodate taxonomy differences

Content is routed to a specified location based on the criteria above

# Acronym Party!

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing

# The Technologies: OCR

**OCR – Optical Character Recognition**

ICR – Intelligent Character Recognition

OMR – Optical Mark Recognition

IDR – Intelligent Document Recognition

Barcode

Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition

Assisted Capture

Fixed Form Process

Semi-Structured Forms Processing

Unstructured Document Processing

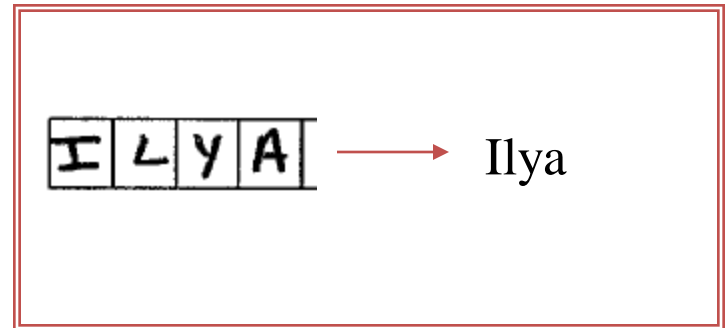
**Ship To:** → Ship To:

# The Technologies: ICR

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing



# The Technologies: OMR

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
**OMR – Optical Mark Recognition**  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing

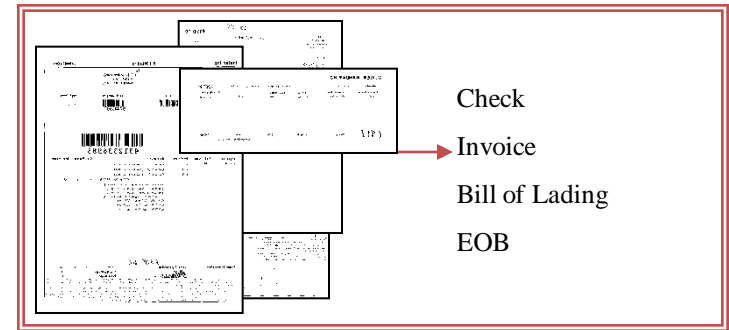
Settlement Account	<input type="checkbox"/>	→	Card Account
Card Account	<input checked="" type="checkbox"/>		

# The Technologies: IDR

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
**IDR – Intelligent Document Recognition**  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing



# The Technologies: Barcode

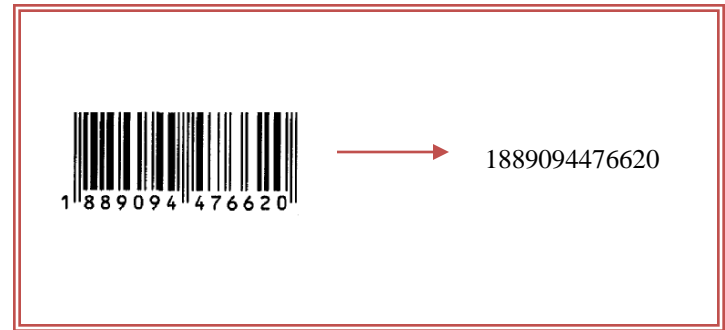
OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition

## Barcode

Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing



# The Technologies: Handwriting

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing

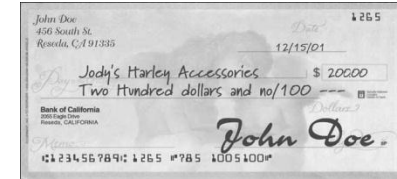
\* CRITICAL \* → \* Critical \*

# The Technologies: CAR/LAR

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



**CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition**  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing



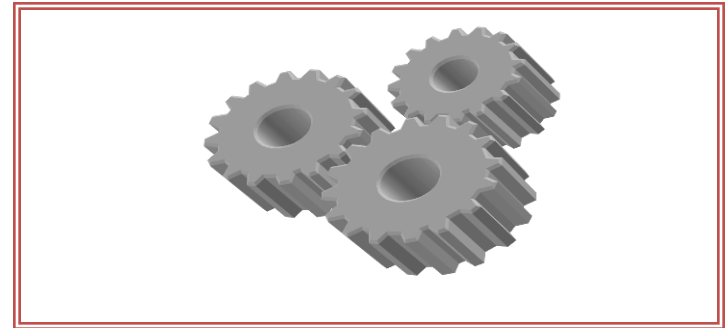
2 hundred dollars & no cents

# The Technologies: Assisted Capture

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
**Assisted Capture**  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing

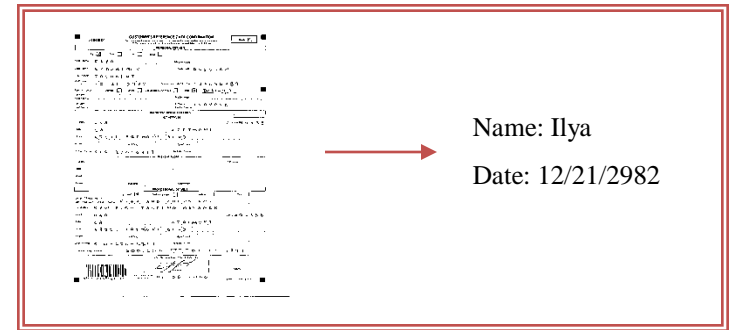


# The Technologies: Fixed Form Processing

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
**Fixed Form Process**  
Semi-Structured Forms Processing  
Unstructured Document Processing



# The Technologies: Fixed Form Processing

**FORM 01** **CUSTOMER'S REFERENCE DATA CONFIRMATION**  Main

Bank guarantees the accuracy of information on the Customers whose specified names in accordance with the General Terms and Conditions and the procedures established by applicable laws

**PERSONAL DETAILS**

Mr  Mrs  Ms  Miss

First Name: ILYIA Middle Initial:

Last Name: EVDOKIMOV Nationality: RUSSIAN

City of Birth: TASHKENT

Birth Date (mm-dd-yyyy): 12/21/2982 Social Security Number: 123456789

Marital Status: Married  Single  Separated/Divorced  Other  Dating

Spouse First Name:  Middle Initial:

Spouse Last Name:  Mother's Maiden Name: IVANOVA

**ADDRESS/CONTACT DETAILS**

**FIXED ADDRESS**

Country: USA ZIP code: 94538

State: CA City: FREMONT

Street: 47221 FREMONT BLVD

House:  Building:  Apartment:

Home Phone: 510-226-6717 Mobile Phone:

**PRESENT ADDRESS**

Country:  ZIP code:

State:  City:

Street:

House:  Building:  Apartment:

**PROFESSIONAL DETAILS**

Employed  Self-employed  Retired  Other

Name of Employer, or if Self-Employed, Company Name: FISH AND CHIPS CO.

Occupation: RAW FISH TASTING MANAGER

Country: USA ZIP code: 94538

State: CA City: FREMONT


Street: 47221 FREMONT BLVD

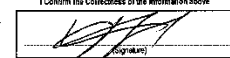
House:  Building:  Apartment:


Work Phone: 510-226-6717 Mobile Phone:

Gross Annual Income: 1500.00 USD Employment Date (mm-dd-yyyy): 01/01/1999

I Confirm the Correctness of the Information above

 122095161180

 Document Date (mm-dd-yyyy): 01/06/2006

 Stamp (Employer HR Officer Signature)



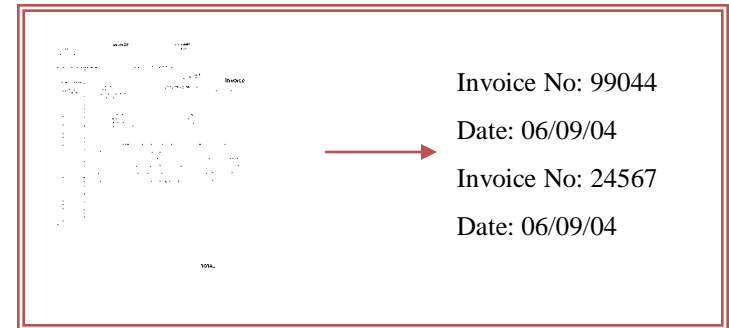
Name: Ilya  
Date: 12/21/2982

# The Technologies: Semi-Structured Forms

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
**Semi-Structured Forms Processing – Complexity is Underestimated**  
Unstructured Document Processing



# The Technologies: Semi-Structured Forms

Note, many people confuse these documents as fixed

206185

**Invoice**

AA APPAREL  
1 Memphis Dr.  
Memphis, TN, 38118  
123-456-7890

RECEIVED MAR - 3

DATE	INVOICE NO.
6/09/04	99044

BILL TO	SHIP TO
US NORTHLAND Doug Darrell 18 Prince George Drive Prince George, VA 23875	US NORTHLAND Doug Darrell 18 Prince George Drive Prince George, VA 23875

P.O. NO.	TERMS	REP	SHIP DATE	SHIP VIA	FOB	PROJECT
1505	Net 30		6/08/04	Courier		

QTY	ITEM	DESCRIPTION	RATE	AMOUNT
7		49910 - Gloves or mittens, NOI, in boxes	150.00	1,050.00
42		15560 - Bulk, NOI, inflated	275.00	11,550.00
24		15520 - Athletic or Sporting Goods, NOI	317.00	7,608.00
<b>TOTAL</b>				20,208.00

INVOICE NUMBER  
24567

CUSTOMER NUMBER  
0598C

PAGE: 1

SHIP TO: CUTLER - WANDAR CANADA  
150 MAIN STREET  
TONAWANDA NY 14150

INVOICE DATE	SLISE	SHIPPING INSTRUCTIONS
6/08/2004	0032	Ground

UNIT PRICE	TAX	PRODUCT TOTAL
10.00		50.00
13.00		130.00
7.00		70.00
12.00		60.00
8.00		80.00
5.00		25.00
11.00		110.00
10.00		50.00
7.00		70.00
8.00		80.00
3.00		15.00
18.00		180.00
7.00		35.00
8.00		80.00
10.00		100.00
11.00		55.00
3.00		30.00
3.00		30.00
10.00		50.00
11.00		55.00
7.00		35.00
11.00		110.00
10.00		100.00
5.00		25.00



Invoice No: 99044  
Date: 06/09/04

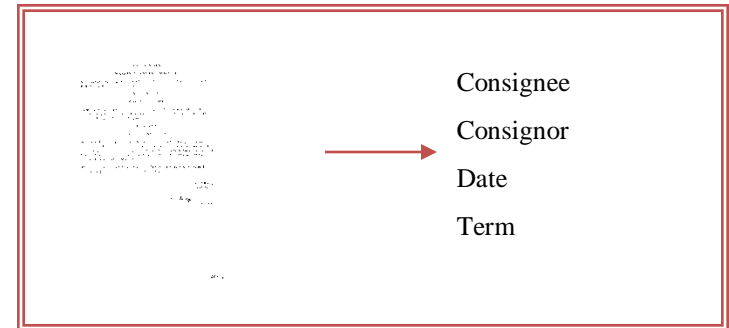
Invoice No: 24567  
Date: 06/09/04 (06/09/2004)

# The Technologies: Semi-Structured Forms

OCR – Optical Character Recognition  
ICR – Intelligent Character Recognition  
OMR – Optical Mark Recognition  
IDR – Intelligent Document Recognition  
Barcode  
Handwriting



CAR/LAR ( Check21 ) – Courtesy and Legal Amount Recognition  
Assisted Capture  
Fixed Form Process  
Semi-Structured Forms Processing  
Unstructured Document Processing



# Common Processes

- Full page conversion
- Classification
- Index level extraction
- Full automation
  
- Redaction
- Routing
- Auto-Filing
- Re-Purposing
- Image Rotation

# The Technologies: Full page conversion

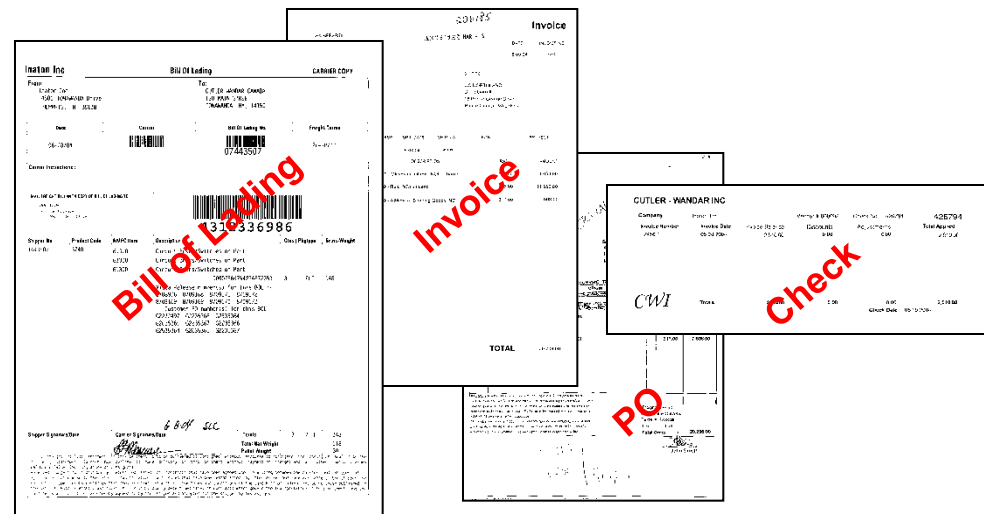
- Image file to electronic data file
- ALL text on the page
- Includes:
  - Image Pre-processing
  - Document Analysis/Zoning
  - Extraction
  - Export ( Commonly PDF, DOC )

# The Technologies: Classification

Software tells you the document type  
Several Modes of document classification

- Image Based
- Contextual

Scan batches of mixed documents





# The Technologies: How Accurate

- Better question is how do you determine accuracy
- Document Type Accuracy
- Field/Zone Location Accuracy
- Data Type Accuracy
- Character Accuracy

# Who are the players?

## Core OCR – ONLY 4!

- ABBYY
- Nuance
- ReadIRIS
- Open Text Document Technologies
- Handful of open source and specialized engines

## SharePoint Ready Solutions

- Psigen
  - KnowledgeLake
  - Kofax
- 
- Everyone Else

# Document Imaging Best Practices

- Setup SharePoint First!
  - ECM: Taxonomy, Content Types, Records Management
  - Platform: RBS, Conversion Services, Content Organizer
- Scan at 300 DPI
- Scan to TIFF Group 4 and convert later
- Start OFF with the basics
- Always test on your documents

- ECM

[www.aiim.org](http://www.aiim.org)

- ECM & SharePoint:

<http://www.aiim.org/Community/SharePoint>

<http://www.aiim.org/Community/Capture>

- Chris's Blog:

<http://blog.livinganalytics.com>

- SharePoint Imaging Toolkit:

<http://www.aiim.org/Shop/Product/2450>

The screenshot shows the AIIM website interface. At the top, the AIIM logo is on the left, and the tagline 'Your Information Management & Collaboration Resource' is in the center. To the right of the tagline are 'Join' and 'Login' buttons. Below the tagline is a search bar with a 'Go' button. A navigation menu contains links for Home, About AIIM, Events, Membership, Community, Training, Research, Resource Center, and My Account. The main content area shows the breadcrumb 'Home > AIIM Shop > Product > User Guide: Document...' and the product title 'User Guide: Document Imaging in SharePoint'. A product image shows the 'AIIM Membership Tools' box set, including a 'User Guide' and 'Document Capture for SharePoint'. The description states: 'This guide is designed to help organizations understand the possibilities of document capture for SharePoint and provide the steps for planning document capture solutions and best practices related to document capture. With proper configuration of document imaging and SharePoint, organizations can ingest paper as conveniently and accurately as electronic data.' To the right is an 'Add to Cart' section with a price table:

Add to Cart	
List Price	\$99.00
Professional Member Price	\$0.00

Below the price table are input fields for 'Username or email' and 'Password', a 'Log In' button, and a 'Stay signed in' checkbox. At the bottom of the login section are links for 'Forgot password?' and 'New user?'.



**Email:** [Chris.Riley@ShareSquared.com](mailto:Chris.Riley@ShareSquared.com)

**Blog:** [Blog.LivingAnalytics.com](http://Blog.LivingAnalytics.com)

**Twitter:** [@RileyBeebs](https://twitter.com/RileyBeebs)

Thank You!

**PLEASE COMPLETE  
THE SESSION SURVEY**