

Course Guide

IBM Datacap 9.0.1: Application Builder

Course code WF261 ERC 1.1



January 2020 edition

Notices

This information was developed for products and services offered in the US.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
United States of America*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

© Copyright International Business Machines Corporation 2017, 2020.

This document may not be reproduced in whole or in part without the prior written permission of IBM.

US Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Trademarks	xiv
Course description	xv
Agenda	xviii
Unit 1. Introduction to Datacap	1-1
Unit Objectives	1-2
Lesson 1.1. Datacap overview	1-3
Datacap overview	1-4
Lessons	1-5
Why is this lesson important to you?	1-6
Content is critical to every business	1-7
The Paper Problem still exists	1-8
Cost of shipping paper adds up	1-9
Business problems that can be solved with Datacap	1-10
Business objectives	1-11
Datacap solution for paper-based business problems	1-12
What is IBM Datacap?	1-13
Structured documents	1-14
Unstructured documents	1-15
Review questions	1-16
Review answers	1-17
Lesson 1.2. Datacap process	1-18
Datacap process	1-19
Lessons	1-20
Why is this lesson important to you?	1-21
IBM Datacap at a Glance	1-22
Source: Distributed Capture	1-23
Datacap supported file formats for the source	1-24
Transform: Recognition, Classification, and Validation	1-25
Automatic Recognition from Images	1-26
Automatic Data Extraction	1-27
Datacap methods for page identification	1-28
Page identification – Barcodes	1-29
Page identification – Keywords	1-30
Page identification – Fingerprints	1-31
Page identification – Pattern Recognition	1-32
Page identification – Text analytics	1-33
Deliver: Captured Documents and Data	1-34
Datacap Capture process	1-35
Rules and rulesets	1-36
Data verification	1-37
Review questions (1)	1-38
Review answers (1)	1-39
Review questions (2)	1-40
Review answers (2)	1-41
Lesson 1.3. Role-based Datacap clients	1-42
Role-based Datacap clients	1-43
Lessons	1-44

Why is this lesson important to you?	1-45
Role-based Datacap tools	1-46
Datacap Navigator – Web client	1-48
Datacap Mobile Capture	1-49
IBM Datacap Mobile app	1-51
Datacap Desktop – Windows based client	1-52
Datacap FastDoc – Windows based client	1-53
Datacap Applications	1-54
Datacap Accounts Payable	1-55
Datacap Medical Claims	1-56
Role-based Datacap clients	1-57
Review questions	1-58
Review answers	1-59
Lesson 1.4. Architecture configurations.	1-60
Architecture configurations	1-61
Lessons	1-62
Why is this lesson important to you?	1-63
Architecture configurations	1-64
Architecture configurations	1-66
Centralized configuration	1-67
Distributed configuration	1-68
Datacap Three Tier Architecture	1-69
Mobile configuration	1-71
Review questions (1)	1-72
Review answers (1)	1-73
Review questions (2)	1-74
Review answers (2)	1-75
Review questions (3)	1-76
Review answers (3)	1-77
Lesson 1.5. Architecture components	1-78
Architecture components	1-79
Lessons	1-80
Why is this lesson important to you?	1-81
Datacap components	1-82
Components - Datacap Navigator and Datacap Web Client	1-84
Components – Datacap Client Workstation	1-86
Components – Developer Workstation	1-87
Components – Datacap Navigator Server	1-89
Components – Datacap Web Server	1-90
Components – Datacap Web Services	1-92
Components – Datacap Server	1-93
Components – Datacap Database Server	1-95
Components – Maintenance Manager Server	1-97
Components – Rulerunner Server	1-99
Components – File Server	1-101
Components – External Systems	1-102
Datacap Folders for Services and Client Code	1-104
Other Datacap Folders	1-106
The Datacap.xml file	1-107
Review question (1)	1-108
Review answers (1)	1-109
Lesson 1.6. Datacap Desktop	1-110
Datacap Desktop	1-111
Lessons	1-112
Why is this lesson important to you?	1-113
Capabilities Overview	1-114

Use Datacap Desktop with your Datacap applications	1-115
Datacap Desktop – Select applications and tasks	1-117
Datacap Desktop - Monitor view	1-118
Select columns and apply filter in the Monitor view	1-119
Exercise: Explore Datacap Desktop interface	1-120
Exercise objectives	1-121
Lesson 1.7. Application design	1-122
Application design	1-123
Lessons	1-124
Why is this lesson important to you?	1-125
Scenario for the complete application	1-126
Steps to design and configure an application	1-127
Do a complete end-to-end evaluation	1-128
Analyze and define the requirements	1-129
Analyze the input documents	1-130
Implement a sandbox system	1-131
Configure Datacap Application	1-132
Application development strategy	1-133
Application Wizard mode	1-135
Application Wizard New Application window	1-137
Test and adjust	1-139
Best Practices	1-140
IBM Advanced Document Imaging PIE	1-141
Exercise: Scan and process a batch	1-142
Exercise objectives	1-143
Lesson 1.8. Introduction to Datacap Navigator	1-144
Introduction to Datacap Navigator	1-145
Lessons	1-146
Why is this lesson important to you?	1-147
What is Datacap Navigator?	1-148
Views in Datacap Navigator	1-149
Datacap View for the business users	1-150
Job Monitor	1-152
Scan Batches	1-153
Rescan options in the Scan tab	1-154
Upload Batches, Page ID, and Profiler tasks	1-155
Classify Batches	1-156
Verify batches	1-157
User settings	1-159
Accessibility features of Datacap Navigator interface	1-160
Exercise: Explore the Datacap Navigator interface and process a batch	1-162
Exercise objectives	1-163
Lesson 1.9. Datacap web client (tmweb)	1-164
Datacap web client (tmweb)	1-165
Lessons	1-166
Why is this lesson important to you?	1-167
Datacap Web Client	1-168
Operations – Task Processing	1-170
Monitor – Job and Task Progress and Statistics	1-171
Administrator – Configuration	1-173
Workflows – Job and Task Definition	1-174
Workflows – Task Setup	1-176
Add Users and Groups	1-178
Add Users to Groups and Add Stations	1-179
Set Privileges	1-181
Set Permissions	1-182

Configure Shortcuts	1-183
Review questions (1)	1-184
Review answers (1)	1-185
Review questions (2)	1-186
Review answers (2)	1-187
Exercise: Datacap Web Client	1-188
Exercise objectives	1-189
Unit summary	1-190
Unit 2. FastDoc Local Mode.....	2-1
Unit Objectives	2-2
Lesson 2.1. FastDoc Overview	2-3
FastDoc Overview	2-4
Lessons	2-5
Why is this lesson important to you?	2-6
What Is FastDoc?	2-7
FastDoc Versatility	2-8
FastDoc interfaces	2-9
FastDoc Interface Icons	2-10
FastDoc overview (1)	2-11
FastDoc overview (2)	2-13
FastDoc overview (3)	2-14
FastDoc Terms (1)	2-15
FastDoc Terms (2)	2-16
Lesson 2.2. FastDoc local mode	2-19
FastDoc local mode	2-20
Lessons	2-21
Why is this lesson important to you?	2-22
FastDoc Local Batch Configuration	2-23
Configure Local Batch Profiles	2-24
Configure the Batch Parameters	2-25
Configuration Local Tasks	2-27
Two FastDoc workstation configurations	2-29
Create and run local batches	2-31
Exercise: FastDoc local mode	2-32
Exercise objectives	2-33
Lesson 2.3. Create a local capture batch	2-34
Create a local capture batch	2-35
Lessons	2-36
Why is this lesson important to you?	2-37
Define terms	2-38
What is a Local Batch Profile?	2-39
What is a Local Task?	2-40
When is a Local Batch Profile required?	2-41
Where is data located?	2-42
Configure local batches	2-44
FastDoc (Admin) local configuration interface	2-46
Batch task pane	2-48
Scan Interface	2-50
Scan Panels	2-52
Exercise: Create a local capture batch	2-53
Exercise objectives	2-54
Lesson 2.4. Process local batches	2-55
Process local batches	2-56
Lessons	2-57
Why is this lesson important to you?	2-58

Scan batch preparation	2-59
Start a batch in local mode	2-60
Scan the batch	2-61
Process the batch	2-63
Verify the batch	2-65
Correct errors	2-67
Export the batch	2-69
Exercise: Process local batches	2-71
Exercise objectives	2-72
Lesson 2.5. Configure Scan and Upload	2-73
Configure Scan and Upload	2-74
Lessons	2-75
Why is this lesson important to you?	2-76
Add an Upload Batch	2-77
Batch profile tasks configuration options	2-78
Configure the Upload Task	2-80
Exercise: Configure scan and upload	2-82
Exercise objectives	2-83
Unit summary	2-84
Unit 3. FastDoc Datacap Server mode	3-1
Unit Objectives	3-2
Lesson 3.1. Introduction to FastDoc Datacap Server mode	3-3
Introduction to FastDoc Datacap Server mode	3-4
Lessons	3-5
Why is this lesson important to you?	3-6
FastDoc Datacap Server interface	3-7
Introduction to FastDoc Datacap Server mode	3-8
Scan Task Interface	3-9
Configure document, pages and fields view	3-11
Configure Ruleset Properties	3-12
Configure workflow view	3-14
Configure Fingerprints	3-16
Run Batches in Datacap Server Mode	3-17
Exercise: FastDoc Datacap Server mode	3-19
Exercise objectives	3-20
Lesson 3.2. Create an application with Form template	3-21
Create an application with Form template	3-22
Lessons	3-23
Why is this lesson important to you?	3-24
What is a FastDoc application template	3-25
Template Folder Structure	3-27
Create an Application	3-28
Wizard Mode window	3-29
Create a FastDoc Application	3-30
Create an application with Form template	3-33
Exercise: Create an application with Form template	3-35
Exercise objectives	3-36
Lesson 3.3. Create batch structure	3-37
Create batch structure	3-38
Lessons	3-39
Why is this lesson important to you?	3-40
Datacap Server mode	3-41
Start FastDoc and login to Datacap Server mode	3-42
Configure a document	3-43
Configure a page	3-44

Document integrity	3-45
Add Fields	3-46
Field Settings	3-47
Exercise: Configure batch structure	3-48
Exercise objectives	3-49
Lesson 3.4. Configure form template rulesets.....	3-50
Configure form template rulesets	3-51
Lessons	3-52
Why is this lesson important to you?	3-53
Configure Forms template rulesets	3-54
Import files - Batch level ruleset	3-56
Convert Files To Images – Batch level ruleset	3-58
Image enhancement – Batch level ruleset	3-59
Image Enhancement ruleset	3-61
Recognize Pages and Fields ruleset (page object)	3-62
Fields properties settings	3-63
Fields properties settings	3-65
Recognize Pages and Fields ruleset (field object)	3-67
Exercise: Configure form template rulesets	3-69
Exercise objectives	3-70
Lesson 3.5. Define fingerprints and test rulesets.....	3-71
Define fingerprints and test rulesets	3-72
Lessons	3-73
Why is this lesson important to you?	3-74
Fingerprint - Define Fingerprint	3-75
Fingerprint – Select Zones	3-76
What is design time testing?	3-77
Test pane profiles and rules	3-79
Test page history and results	3-80
Test pane error resolution	3-83
Exercise: Define fingerprints and test rulesets	3-84
Exercise objectives	3-85
Lesson 3.6. Populate fields with keyword lookup	3-86
Populate fields with keyword lookup	3-87
Lessons	3-88
Why is this lesson important to you?	3-89
Add the Locate Fields Using Keyword ruleset	3-90
Configure Populate Fields Using Keyword ruleset	3-91
Select population methods	3-92
Locate the lookup term	3-93
Locate value with direction and grouping	3-95
Exercise: Populate fields with keyword lookup	3-97
Exercise objectives	3-98
Lesson 3.7. Configure validate field rulesets.....	3-99
Configure validate field rulesets	3-100
Lessons	3-101
Why is this lesson important to you?	3-102
Validate Fields ruleset	3-103
Common ruleset options	3-104
Length & Value Content ruleset options	3-106
Date ruleset options	3-107
Currency & Numeric ruleset options	3-108
Lookup ruleset options	3-110
Custom Error Message ruleset options	3-111
Exercise: Configure validate fields ruleset	3-112
Exercise objectives	3-113

Lesson 3.8. Export to IBM FileNet Content Manager	3-114
Export to IBM FileNet Content Manager	3-115
Lessons	3-116
Why is this lesson important to you?	3-117
Delivery options	3-118
Export configuration options	3-120
General export action	3-122
Select a Ruleset	3-123
Configure the Batch parameters	3-124
Configure the Document parameters	3-125
Configure the Field parameters	3-127
View IBM Repositories with Datacap Navigator	3-128
IBM FileNet repository configuration parameters	3-130
View Exported objects in the repository	3-131
Exercise: Export to IBM FileNet Content Manager	3-132
Exercise objectives	3-133
Unit summary	3-134
Unit 4. Datacap Studio Introduction	4-1
Unit Objectives	4-2
Lesson 4.1. Datacap Studio – Rulemanager	4-3
Datacap Studio – Rulemanager	4-4
Lessons	4-5
Why is this lesson important to you?	4-6
Panel Organization within Datacap Studio	4-7
Datacap Studio Overview	4-10
Rulemanager View	4-12
Document Hierarchy (DCO)	4-14
Rules and Rulesets	4-16
Rules and Rulesets	4-17
Actions and Functions	4-19
Global and Application Rulesets	4-20
Ruleset Types	4-22
Rule Mapping	4-24
Properties tab	4-26
Ruleset object properties tab	4-28
Rule object properties tab	4-29
Function and Action object properties tab	4-31
Action object properties tab	4-32
Task Profiles	4-33
Actions Library	4-35
Exercise: Explore Datacap Studio - Rulemanager	4-36
Exercise objectives	4-37
Lesson 4.2. Datacap Studio – Zones	4-38
Datacap Studio – Zones	4-39
Lessons	4-40
Why is this lesson important to you?	4-41
Zones View	4-42
Fingerprints View	4-44
Document Hierarchy View	4-45
Properties View	4-46
Image View	4-47
Exercise: Datacap Studio Zones	4-48
Exercise objectives	4-49
Unit summary	4-50

Unit 5. Add a Multi Page Document	5-1
Unit Objectives	5-2
Lesson 5.1. Create a Multi-Page Form Application	5-3
Create a Multi-Page Form Application	5-4
Topics	5-5
Why is this lesson important to you?	5-6
Image Samples	5-7
Application template choice	5-8
Configure document, pages and fields view	5-9
Add an HRF_Document and HRF_Page	5-10
HRF_Page Settings	5-12
Recognition considerations	5-13
Full page zonal recognition	5-15
Recognize hand written text fields (ICR)	5-16
Recognize mark zones (OMR)	5-17
Field Settings tab	5-19
Configure Fingerprints	5-21
Add a Donation_Receipt Page	5-22
Exercise: Create a Multi-Page Form Application	5-23
Exercise introduction	5-24
Lesson 5.2. Recognize with ICR and OMR	5-25
Recognize with ICR and OMR	5-26
Topics	5-27
Why is this lesson important to you?	5-28
Overview	5-29
Recognize with the ICR Engine	5-30
More Recognize Actions	5-31
Datacap Studio Zones tab	5-32
ICR/C Tab Settings	5-34
Recognize OMR	5-36
Establishing OMR Parent fields	5-37
Select OMR Properties View	5-38
Set OMR Length Property	5-39
The OMR Settings	5-40
Exercise: Recognize with ICR and OMR	5-42
Exercise introduction	5-43
Lesson 5.3. Validate with database lookup	5-44
Validate with database lookup	5-45
Topics	5-46
Why is this lesson important to you?	5-47
Steps to Configure Lookup	5-48
Configure a lookup database connection string	5-50
Database Lookup Actions	5-51
Validate Fields Zip Lookup Function	5-52
Exercise: Validate with database lookup	5-53
Exercise introduction	5-54
Unit summary	5-55
Unit 6. Basic Learning Application	6-1
Unit Objectives	6-2
Lesson 6.1. Create a Learning Template application	6-3
Create a Learning Template application	6-4
Lessons	6-5
Why is this lesson important to you?	6-6
Create a Learning Template-based application	6-7
Characteristics of a Learning Template application	6-9

Characteristics of a Learning Template DocType	6-10
Learning Template configuration	6-11
Main_Page	6-12
Document integrity Values	6-13
Document integrity	6-14
Learning Template fields (1)	6-15
Learning Template fields (2)	6-16
Exercise: Create a Learning Template application	6-17
Exercise objectives	6-18
Lesson 6.2. Configure a Learning Template application	6-19
Configure a Learning Template application	6-20
Lessons	6-21
Why is this lesson important to you?	6-22
The Learning Template and fingerprints	6-23
The Learning Template workflow	6-24
More about the Learning Template workflow	6-26
More about the Learning Template workflow	6-27
More about the Learning Template	6-28
What about routing and export?	6-29
More things to know	6-30
Exercise: Configure a Learning Template application	6-31
Exercise objectives	6-32
Lesson 6.3. Create locate rules	6-33
Create locate rules	6-34
Lessons	6-35
Why is this lesson important to you?	6-36
The basic locate actions	6-37
Rule configuration and run explanation	6-38
Locate actions (1)	6-39
Locate actions (2)	6-40
Locate and zone actions	6-41
Rental_Agreement page	6-42
Locate rule for customer name field	6-43
Locate rule for pickup date	6-44
Locate rule for agreement number field	6-45
Locate rule for total field	6-46
Exercise: Create locate rules	6-48
Exercise objectives	6-49
Lesson 6.4. Create validate rules	6-50
Create validate rules	6-51
Lessons	6-52
Why is this lesson important to you?	6-53
The validate ruleset	6-54
Validate and export task relationship	6-56
Sample Validation actions	6-57
Assign an action to the Field Name rule	6-58
Validations actions	6-59
Validations actions	6-60
Mixing locate and validate actions	6-61
Exercise: Create validate rules	6-62
Exercise objectives	6-63
Lesson 6.5. Expand the locate rules	6-64
Expand the locate rules	6-65
Lessons	6-66
Why is this lesson important to you?	6-67
Locate rule for passenger name field	6-68

Locate rule for the ticket number field	6-69
Locate rule for issue dates	6-70
Exercise: Expand the locate rules	6-71
Exercise objectives	6-72
Lesson 6.6. Expand the validate rules	6-73
Expand the validate rules	6-74
Lessons	6-75
Why is this lesson important to you?	6-76
Validate rule for the issue dates	6-77
Validate rule for the ticket number field	6-78
Verify	6-79
Exercise: Expand the validate rules	6-80
Exercise objectives	6-81
Lesson 6.7. Export to an IBM FileNet repository	6-82
Export to an IBM FileNet repository	6-83
Lessons	6-84
Why is this lesson important to you?	6-85
Export types	6-86
Preparation for export to IBM FileNet repository	6-87
Actions to export to IBM FileNet repository	6-88
FastDoc batch level configuration	6-89
Function to connect to the IBM FileNet server	6-90
FastDoc document level configuration	6-91
FastDoc field level configuration	6-92
Setting the value of a document's properties	6-93
Configure multi-page documents	6-94
Separator page sample	6-95
Exercise: Export to an IBM FileNet repository	6-96
Exercise objectives	6-97
Unit summary	6-98
Unit 7. Debug and Test	7-1
Unit Objectives	7-2
Lesson 7.1. Application Debugging	7-3
Application Debugging	7-4
Topics	7-5
Why is this lesson important to you?	7-6
Application Debugging and Datacap Studio	7-7
The Logs Used in Debugging	7-8
The RRS Log Files	7-9
Web Client and Datacap Desktop Logging	7-10
Setting Rulerunner Logging by Application/Task	7-11
Datacap Web Client Tasks and Log Files	7-13
Support Recommendations	7-14
Further Recommendations	7-15
Another Key Way to Debug	7-16
Exercise: Application Debugging	7-17
Exercise objectives	7-18
Lesson 7.2. Datacap Studio Test Tab	7-19
Datacap Studio Test Tab	7-20
Topics	7-21
Why is this lesson important to you?	7-22
Datacap Studio Test Tab and Debugging	7-23
Datacap Studio Test Tab	7-24
Four of the Eight Panels	7-25
Other Four Panels	7-26

Test Tab Debugging Features	7-27
Setting and Using Breakpoints	7-28
Setting Generic Breakpoints	7-29
Breakpoint Configuration Interfaces	7-30
Setting Breakpoints	7-31
Disabling and Clearing Breakpoints	7-32
Single Stepping Through Your Code	7-33
How to Step Through Code: Step in	7-34
How to Step Through Code: Step Over or Out	7-35
Examining Log Files from the Test Tab	7-36
Exercise: Datacap Studio Test tab	7-37
Exercise objectives	7-38
Unit summary	7-39

Trademarks

The reader should recognize that the following terms, which appear in the content of this training document, are official trademarks of IBM or other companies:

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide.

The following are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide:

DB2®

FileNet®

Tivoli®

WebSphere®

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other product and service names might be trademarks of IBM or other companies.

Course description

IBM Datacap 9.0.1: Application Builder

Duration: 5 days

Overview

This comprehensive course covers various aspects of Datacap application building using the FastDoc and Datacap Studio tools. It introduces the interface, and shows how to create and enhance applications from Form and Learning Templates for processing batches.

Audience

- Application Builders who are responsible for building Datacap applications
- Anyone who needs to know how to rapidly develop a capture solution and enhance the application using the full capability of the IBM Datacap suite.

Prerequisites

- Recommended: Familiarity with data capturing concepts

Course Objectives

Upon completion of this course, participants will be able to:

- Work with FastDoc Local mode
 - Use FastDoc as a development tool
 - Configure and run tasks in the local mode
 - Configure a Scan and Upload Task
- Configure applications in FastDoc Server mode
 - Build a new Application with the Form template
 - Create a batch structure and configure form template rulesets
 - Add a page fingerprint
 - Design time testing of ruleset
 - Populate fields with keyword lookup
 - Configure validate field rulesets
 - Configure a new page fingerprint
 - Configure export to rules for exporting to IBM FileNet Content Manager Repository

- Create a multi-page application based on Form Template
 - Use Page Separator pages to configure a multi-page batch
 - Recognize hand written fields with Intelligent Character Recognition ICR
 - Recognize handmade digital marks with Optical Mark Recognition OMR
 - Recognize machine print fields with Optical Character Recognition OCR
 - Validate using a lookup Database
- Work with Datacap Studio
 - Explore Rulemanager, Zones, and Test views in Datacap Studio
- Create an application and configure in FastDoc
 - Create an application based on the Learning Template
 - Configure Batch Structure (Document Hierarchy)
 - Setup Image Enhancement
 - Configure Page recognition
- Enhance the application in Datacap Studio
 - Create the Locate rules to locate fields with keyword search
 - Locate fields with intellocate (automatic in the learning template)
 - Configure the field Validation rules to validate the fields on the pages
 - Extend the Locate and Validation rules to handle more page types
- Export to IBM FileNet Content Manager repository
 - Convert document images to pdf format
 - Configure Export Rules for IBM FileNet Content Manager
- Application debugging and Test the task profiles
 - Use the Test tab in Datacap Studio for application debugging
 - Locate and use Datacap Logs to troubleshoot Datacap error conditions

Course outline

- FastDoc Overview
- FastDoc local mode and Server mode
- Create a local capture batch
- Configure batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate fields ruleset
- Export to IBM FileNet Content Manager
- Configure a multi-page document in a Form Template application
- Data recognition
- Validate with database lookup

- Introduction to Datacap Studio
 - Create an application based on Learning Template
 - Enhance the application in Datacap Studio
 - Create Locate and Zone rules
 - Create Validate rules
 - Expand the Locate and Validation rules
 - Document export to corporate data repository
 - Application Debugging and Logs
-

Agenda



Note

The following lesson durations are estimates, and might not reflect every class experience.

Day 1

(00:15) Welcome

Unit 1. Datacap Introduction

(00:30) Lesson 1.1 Datacap overview

(00:30) Lesson 1.2 Datacap process

(00:30) Lesson 1.3 Role-based Datacap clients

(00:30) Lesson 1.4 Architecture configurations

(01:00) Lesson 1.5 Architecture components

(01:00) Lesson 1.6 Datacap Desktop

(01:00) Lesson 1.7 Application Design

(01:00) Lesson 1.8 Datacap Navigator Introduction

(01:00) Lesson 1.9 Datacap Web Client (tmweb) (optional)

Day 2

Unit 2. FastDoc Local Mode

(00:30) Lesson 1 - FastDoc overview

(00:30) Lesson 2 - FastDoc Local mode

(01:00) Lesson 3 - Create local capture batch

(01:00) Lesson 4 - Process Local batches

(01:00) Lesson 5 - Configure scan and upload

Unit 3. FastDoc Server Mode

(01:00) Lesson 1 - Introduction to FastDoc Datacap Server mode

(01:00) Lesson 2 - Create an Application with Form template

(01:00) Lesson 3 - Create batch structure

Day 3

Unit 3. FastDoc Server Mode

- (01:00) Lesson 4 - Configure form template rulesets
- (01:30) Lesson 5 - Define fingerprints and test rulesets
- (01:00) Lesson 6 - Populate fields with keyword lookup
- (01:00) Lesson 7 - Configure validate field rulesets
- (01:30) Lesson 8 - Export to IBM FileNet Content Manager

Unit 4. Datacap Studio Introduction

- (00:30) Lesson 1 - Datacap Studio - Rulemanager
- (00:30) Lesson 2 - Datacap Studio - Zones and Test views

Day 4

Unit 5. FastDoc Multi-Page Document

- (01:30) Lesson 1 - Create a multi-page form application
- (01:15) Lesson 2 - Recognize with ICR and OMR
- (01:15) Lesson 3 - Validate with database lookup

Unit 6. Basic Learning Application

- (01:00) Lesson 1 - Create a Learning Template application
- (01:00) Lesson 2 - Configure a Learning Template application
- (01:15) Lesson 3 - Create locate rules

Day 5

Unit 6. Basic Learning Application

- (01:15) Lesson 4 - Create validate rules
- (01:30) Lesson 5 - Expand the locate rules
- (01:00) Lesson 6 - Expand the validate rules
- (01:15) Lesson 7 - Export to an IBM FileNet repository

Unit 7. Debug and Test

- (01:00) Lesson 1 - Application debugging
- (01:00) Lesson 2 - Datacap Studio Test tab

Unit 1. Introduction to Datacap

Estimated time

07:00 hours

Overview

This unit introduces you to the business solution that IBM Datacap provides, the Datacap process, and the capabilities of Datacap.

This unit provides an overview of Datacap Architecture and Application Design.

You use Datacap Desktop (Windows based client), Datacap Navigator, and the Datacap web client (tmweb) to process a batch of input data.

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

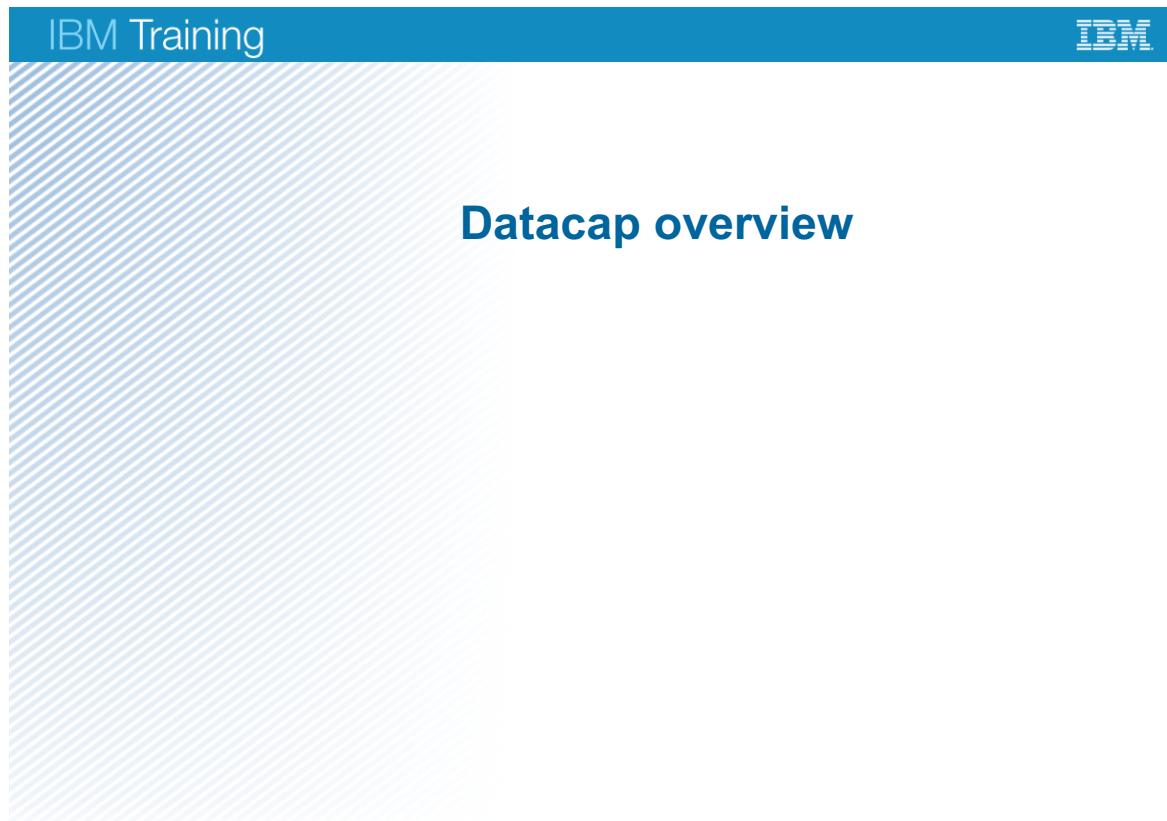
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/datacap_9.0.1.htm

Unit Objectives

- Identify the business solution that IBM Datacap provides, the Datacap process, and the capabilities of Datacap.
- Work with the Datacap Navigator, Datacap Desktop (Windows based client), and the Datacap web client (tmweb) to process a batch of input data.
- Identify the components of Datacap (Architecture) and the things to consider for Application Design.

Figure 1-1. Unit Objectives

Lesson 1.1. Datacap overview



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-2. Datacap overview

Lessons

- ▶ Datacap overview
 - Datacap process
 - Role-based Datacap clients
 - Architecture configurations
 - Architecture components
 - Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-3. Lessons

Why is this lesson important to you?

- This lesson provides an overview of the business solution that IBM Datacap provides, and its capabilities.

Figure 1-4. Why is this lesson important to you?

Content is critical to every business

Every Industry has important business documents that need to be captured.

 Financial Services <ul style="list-style-type: none"> • Account opening • Mortgage applications • Claims processing • Investment reports • Regulations 	 Government <ul style="list-style-type: none"> • Benefits management • Claims • Citizen Correspondence • Social services • Permits and license • Taxes 	 Industrial <ul style="list-style-type: none"> • Standard operating procedures • Engineering document management • Incidents and investigations
 Telecom <ul style="list-style-type: none"> • Cell phone contracts • E-billing and statement processing • Contact center/single view of the customer • Voice of the customer 	 Healthcare <ul style="list-style-type: none"> • Care plans • Critical pathway • Electronic health record • Doctor's notes • Medical claims 	 Horizontal <ul style="list-style-type: none"> • Accounts Payables • Invoice processing • Human resources • Project management • Contract management • Call center

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-5. Content is critical to every business

The Paper Problem still exists



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-6. The Paper Problem still exists

Reasons for still using paper are:

- **Historical**

The existing forms and documents are what the organization must deal with. The organization has no say in the decision of converting these hardcopy forms and documents to electronic formats. Sometimes, the number of existing forms and documents make it unrealistic to do so.

- **Legislation**

Legislation is not keeping pace with new technology and continues to require records to be on paper (which holds probative value).

- **Low-tech portability of paper**

Paper can be transported to customers anywhere, irrespective of such things as affordability, access to infrastructure, technical dependencies, or administrative boundaries. For example, paper can be sent to customers by mail.

Cost of shipping paper adds up

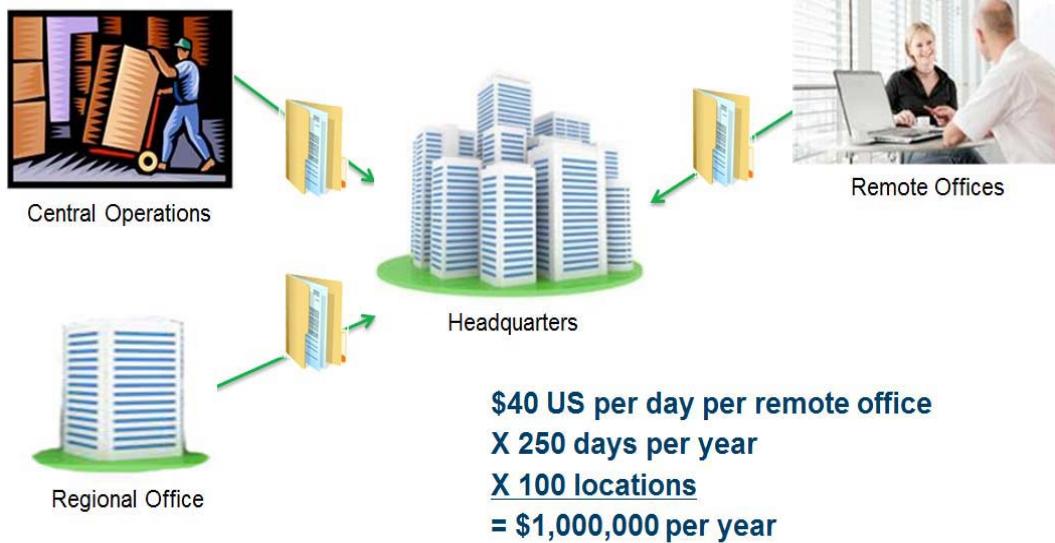


Figure 1-7. Cost of shipping paper adds up

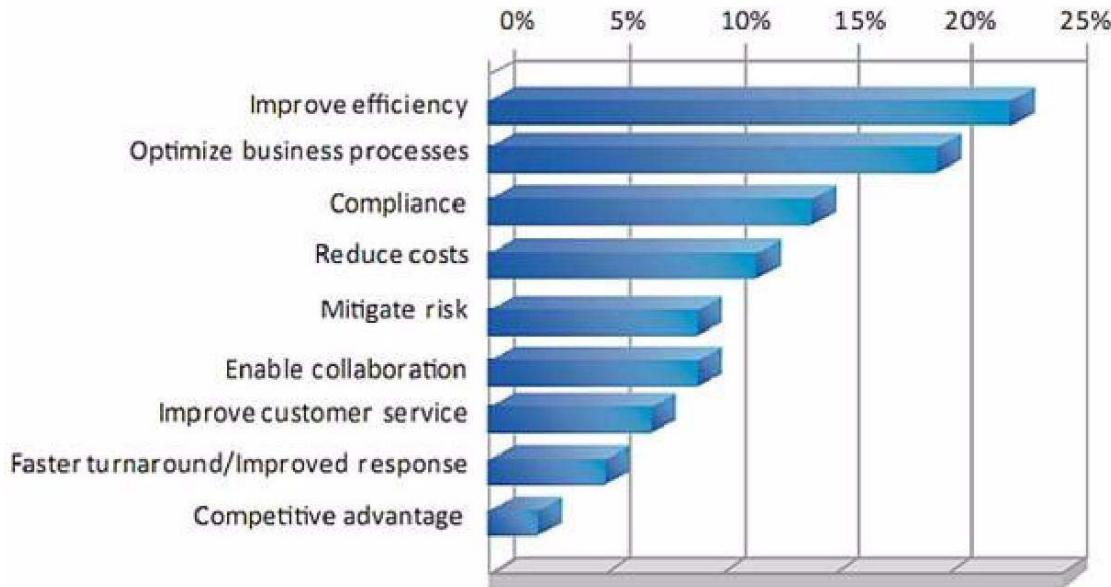
Business problems that can be solved with Datacap

- Datacap provides solution for a list of business challenges that the paper-based documents present:
 - Inefficient, time-consuming, and not flexible.
 - Expensive to ship the paper documents between various business locations.
 - Difficult to store them for the long term.
 - Hard to preserve them optimally for business, legal, disaster (flood and fire), security, and safety reasons.
 - Physical documents can be more easily lost, misfiled, or misclassified and never recovered.
 - Compliance is a concern - preserving the right documents and discarding documents that are no longer needed for the business.
 - Manually extracting the appropriate data from different type of business documents is expensive.

Figure 1-8. Business problems that can be solved with Datacap

IBM Training 

Business objectives



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-9. Business objectives

The challenges that are discussed in previous charts align precisely with the business objectives that many organizations hope to achieve in Enterprise Content Manager systems. These systems include production imaging systems that help solve the paper problems and challenges.

The image shows a recent study that was conducted by the Association for Information and Image Management (AIIM) on business objectives in Enterprise Content Management systems.

Improved efficiency 23%

Optimize business processes 19%

Compliance 14%

Reduce costs 12%

Mitigate risk 9%

Enable collaboration 9%

Improve customer service 7%

Faster Turnaround/Improve response 5%

Competitive advantage 2%

Datacap solution for paper-based business problems

- IBM Datacap Capture can help automate a solution for the enterprise by:
 - Turning paper documents into electronic files.
 - Streamlined process - automating the extraction of appropriate data.
 - Indexing metadata to facilitate searching and quick access.
 - Verifying and saving documents with minimal user intervention.
 - Enabling data sharing across departments, division, and geographic regions.
 - Helping archival and lifecycle management (the scanned documents are exported to repositories).
 - Increasing productivity (accelerating the process and saving time).

Figure 1-10. Datacap solution for paper-based business problems

IBM Datacap was designed for enterprise-wide deployments in paper-intensive market segments, such as government, insurance, healthcare, financial services, and transportation, to name a few.

Increased productivity

By significantly reducing manual data entry and paper-based storage and retrieval of documents, knowledge workers can more quickly process and access documents. Quicker access to documents improves customer service case management, business transactions, and compliance.

Streamlined process

IBM Datacap eliminates a cumbersome paper process by enabling clients to automate previously labor-intensive aspects.

Reduced cost

IBM Datacap is designed to reduce the required data entry personnel by 50 % and sometimes more. Distributed scanning and verification enables clients to reduce or eliminate document shipping costs and distribute labor to areas with more affordable labor rates.

Increased accuracy: By eliminating the errors that tired or distracted human data entry operators historically make, IBM Datacap can save time. Time is saved tracking down and fixing faulty data or misplaced document images.

What is IBM Datacap?

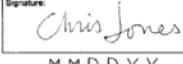
- Advanced Document Imaging and capture software.
- Datacap automates capturing documents.
- Datacap automates the extraction of appropriate data.
 - Automation accelerates the process, and business can respond quickly to their customers and Business Partners.
- Datacap can handle different types of documents regardless of what business processes are involved.
 - Structured documents
 - Unstructured documents

Figure 1-11. What is IBM Datacap?

Structured documents

- All documents have a consistent format.
- Every document has the same data fields in the same place.
- Examples:
 - Tax forms
 - Beneficiary forms
 - Claims
 - Reimbursement forms

Auto Insurance Claim Form
Insurance Company A

Policy Holder Address :	Chris Joes, 13 Johan Street NY 987654
Driver Name :	Chris Joes
Policy Number :	46998813
Incident Number :	CL-4328941
Incident Date:	03/03/11
Incident Time:	14:15
Vehicle License :	2296RG
Vehicle Colour :	Black
Vehicle Manufacturer :	Tos
Vehicle Model :	Quaser
Year of Manufacture :	2009
Chassis Number (VIN) :	2C5BB4CRX82LP3489
Incident Description:	Incident occurred on Highway 69, NY. Damage occurred to rear of the vehicle. Damaged right tail light and bumper.
Was anyone injured in this incident which required medical attention?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
By signing the adjacent box and entering the current date, you agree the information above is accurate to the best of your knowledge. Signature:  M M D D Y Y Dated: 03 SEPTEMBER 2011	

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-12. Structured documents



Unstructured documents

- Documents can have different formats.
- The documents might have the same information, but the same data fields are at different locations in each document.
- Examples:
 - Loan package
 - Contracts
 - Invoices
 - Paystubs

B S		Brilliant Repair Shop 90 Faxed Street Carson City, NV 89701 Phone: 775-555-6789	INVOICE
INVOICE #03-8608 DATE: 02/15/10			
To: Busy Car Repair 100 Auto Road Salt Lake City, UT 84101 Phone: 801-555-1234	Ship To: Busy Car Repair 100 Auto Road Salt Lake City, UT 84101 Phone: 801-555-1234		
Comments or special instructions:			
SALESPERSON 196	PO. NUMBER 0811012	REGISTRATION []	SHIPPED VIA []
TERMS 2/10 Net 30			
QUANTITY	DESCRIPTION	UNIT PRICE	TOTAL
1	Left quarter panel	179.00	179.00
2	Tail lights	50.00	100.00
80	Gaskets	.95	76.00
2	Front headlights	50.00	100.00
80	Bolts	.95	76.00
SUBTOTAL			572.00
SALES TAX			27.68
SHIPPING & HANDLING			
TOTAL DUE			599.68
Make all checks payable to Brilliant Repair Sjop . If you have any questions concerning this invoice, contact (775) 555-6789.			
Thank you for your business!			

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-13. Unstructured documents

Review questions

1. True or False.
Datacap supports both structured and unstructured documents.
2. Which of the following descriptions apply to IBM Datacap? Select more than one option:
 - A. Automates capturing documents and the extraction of appropriate data.
 - B. Improves efficiency and optimizes business processes.
 - C. Supports compliance and Risk mitigation.
 - D. Reduces costs and speeds up the response.
 - E. Improves customer service.

Figure 1-14. Review questions

Review answers

1. True or False.

Datacap supports both structured and unstructured documents.

[The Answer is True.](#)

2. Which of the following descriptions apply to IBM Datacap? Select more than one option:

- A. Automates capturing documents and the extraction of appropriate data.
- B. Improves efficiency and optimizes business processes.
- C. Supports compliance and Risk mitigation.
- D. Reduces costs and speeds up the response.
- E. Improves customer service.

[The Answers are: A, B, C, D, and E](#)

Figure 1-15. Review answers

Lesson 1.2. Datacap process

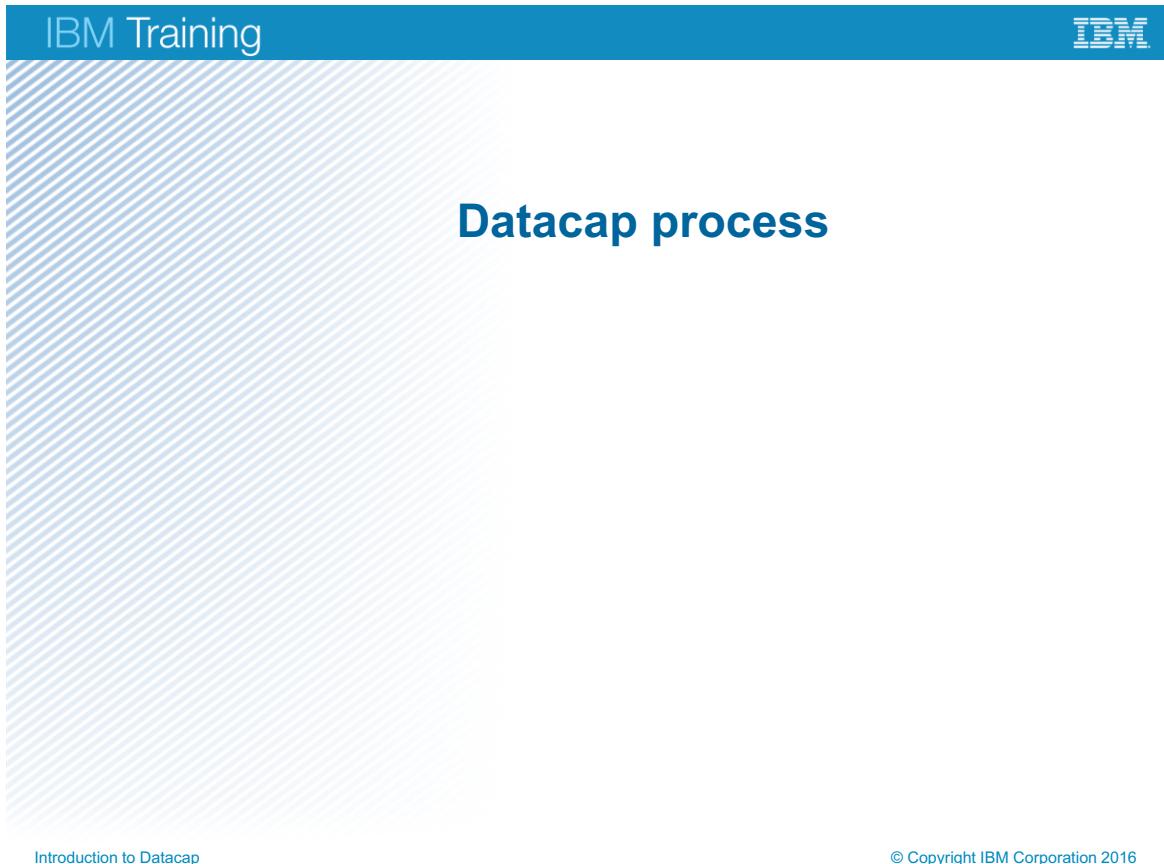


Figure 1-16. Datacap process

Lessons

- Datacap overview
- Datacap process
- Role-based Datacap clients
- Architecture configurations
- Architecture components
- Datacap Desktop
- Application design
- Introduction to Datacap Navigator
- Datacap web client (tmweb)

[Introduction to Datacap](#)

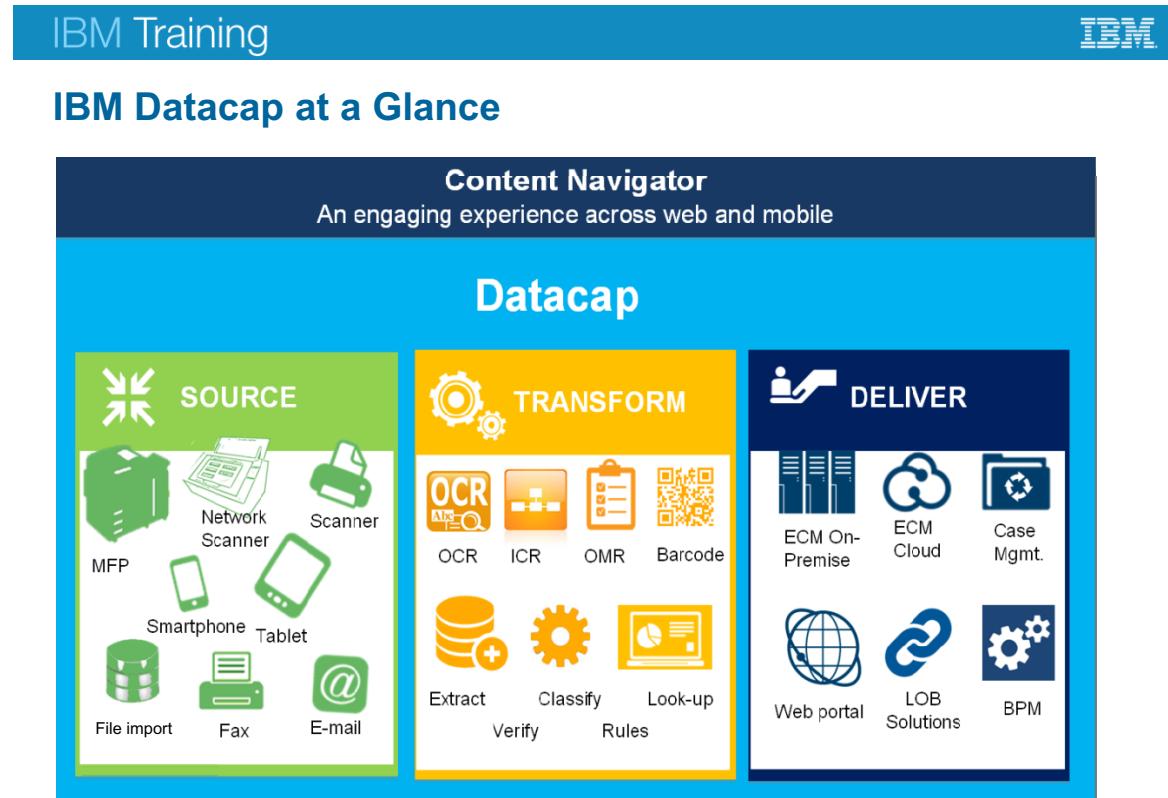
© Copyright IBM Corporation 2016

Figure 1-17. Lessons

Why is this lesson important to you?

- This lesson provides an overview of Datacap process.

Figure 1-18. Why is this lesson important to you?



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-19. IBM Datacap at a Glance

Stages of Datacap process

Source

- Documents that are captured from different channels.

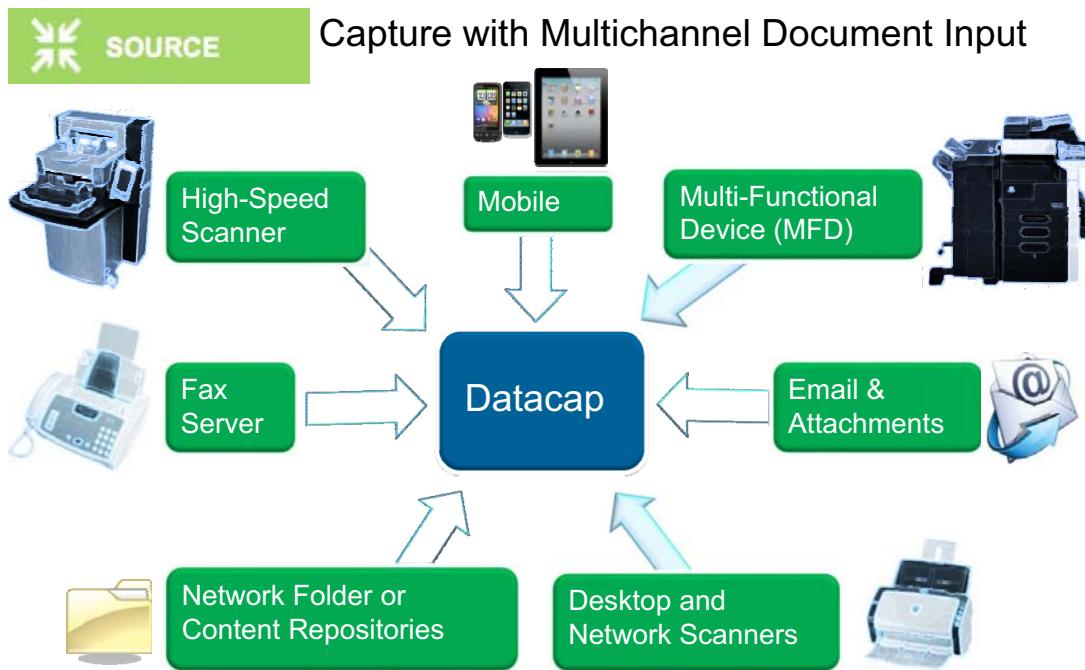
Transform

- Datacap classifies, recognizes, validates, and verifies the document content.

Deliver

- The capture documents can be exported to different back-end systems, and made available for various applications.

Source: Distributed Capture



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-20. Source: Distributed Capture

The diagram shows different channels from which the source documents are captured.

Datacap supported file formats for the source

Input Channels	File Format	Notes
<ul style="list-style-type: none"> • Scanners • Multi-Functional Devices 	<ul style="list-style-type: none"> • TIFF, JPEG, PDF 	<p>Color and grayscale are typically converted to single page bitonal TIFF.</p>
<ul style="list-style-type: none"> • Mobile devices • Email attachments • Windows file system 	<ul style="list-style-type: none"> • TIFF, JPEG, PDF, PNG • HTML, RTF, TXT, DOC, DOCX, XLS, XLSX, ZIP, EML 	<ul style="list-style-type: none"> • JPEG2000 compression is not supported. • Password-protected PDF files and fillable PDF forms are not supported.
<ul style="list-style-type: none"> • FAX 	<ul style="list-style-type: none"> • G3, or G4 TIFF • Single-page or multi-page bitonal 	<ul style="list-style-type: none"> • For recognition and all tasks. • Fax documents are black and white.

Figure 1-21. Datacap supported file formats for the source

For more information, see “Software Product Compatibility Reports Datacap 9.0.1”

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.dc.install.doc/sw_comp_reports.htm

Transform: Recognition, Classification, and Validation



- Machine print and hand print character recognition
- One and two dimensional barcode readings
- Check marks
- Automatic document classification
- Extraction of metadata
- Validation
- Flexible rules engine



Image Technology	Typical Accuracy
Barcode	99+%
OCR	98-99%
ICR	90%
IDR	85-90%

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

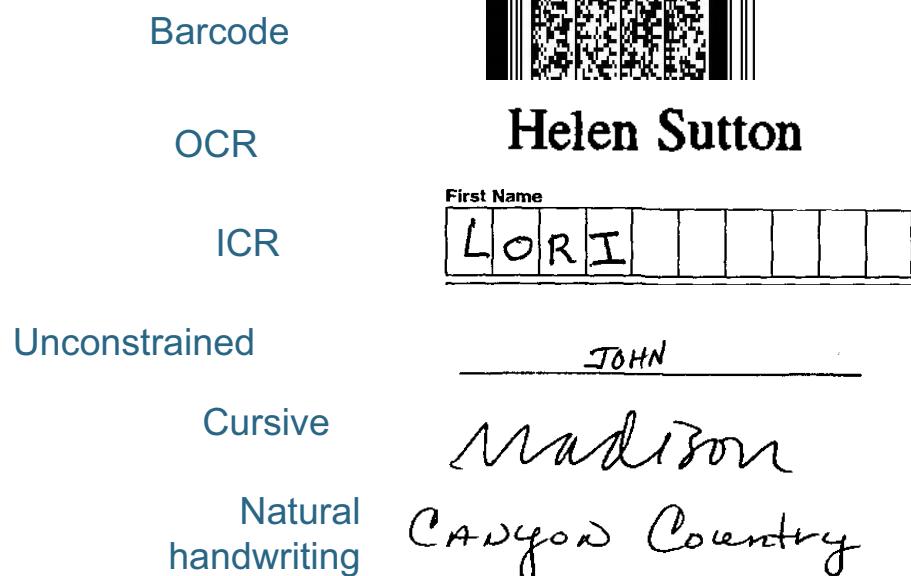
Figure 1-22. Transform: Recognition, Classification, and Validation

By using the Datacap rules engine, data capture can be tailored to fit the most demanding business requirements.

- The settings can be changed quickly when business needs change.

Automatic Recognition from Images

Data Capture Types - Automatic Recognition from Images



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-23. Automatic Recognition from Images

OCR – Optical Character Recognition

- This technology is used to convert system-printed text in an image to editable text

ICR – Intelligent Character Recognition

- Recognize hand written characters

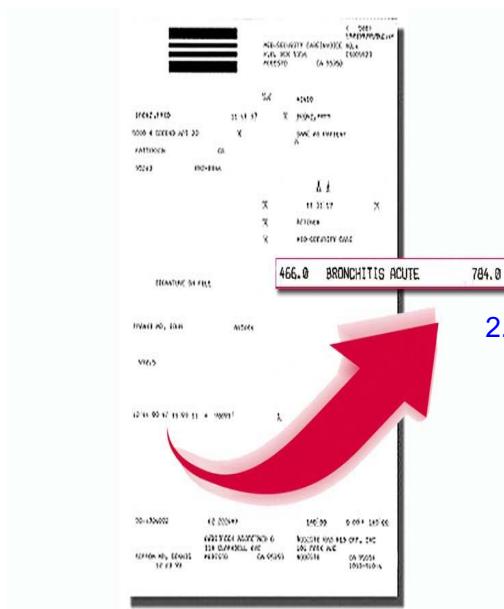
OMR – Optical Mark Recognition

- Recognize check boxes, radio buttons, and so on.

Barcode recognition

- The ability to recognize and interpret barcodes.

Automatic Data Extraction



1. Index the document image with searchable metadata in Enterprise Content Manager
 - Vendor name
 - Invoice number
 - Invoice date
 - SSN
 - Customer number
 - Provider ID
2. Populate Business Systems with Key Data
 - Banking Systems
 - Accounting System
 - ERP
 - CRM
 - Workflow
 - Adjudication
 - Order processing
 - Business Intelligence

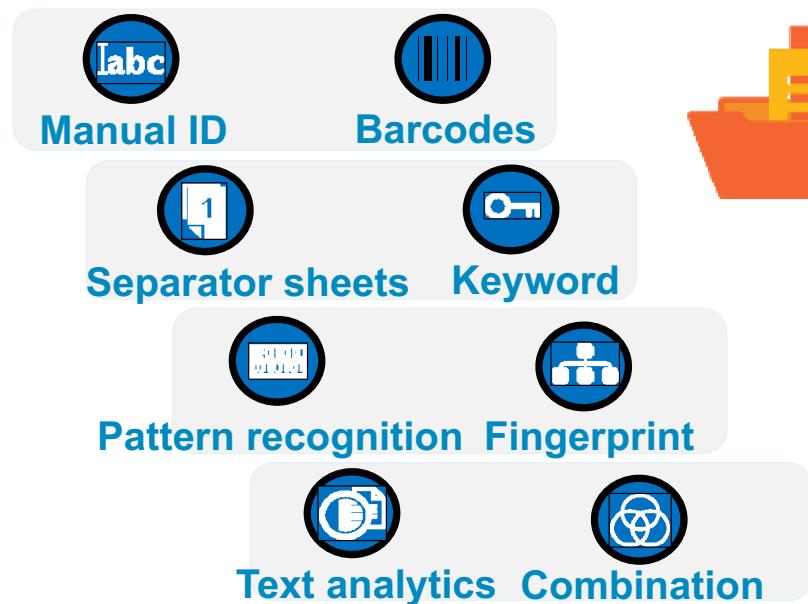
Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-24. Automatic Data Extraction

Datacap methods for page identification

Automatic Classification reduces the workload of manual processing.



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-25. Datacap methods for page identification

The diagram shows various page identification methods.

Document classification is based on the page identification.

Page identification – Barcodes

- Identifies the current page based on the barcode values that are found in the image.
- One- and two-dimensional barcodes are used for page recognition.
- One-dimensional barcodes
 - Can be with or without numbers.
 - Example: Code 39 can store up to 43 alphanumeric characters.
- Two-dimensional (2D) barcodes
 - Coded with a matrix that represents information along the vertical and horizontal axes of the barcode.
 - Can store up to several kilobytes of data.



1234567890



Figure 1-26. Page identification – Barcodes

Page identification – Keywords

- Keyword identifies the current page based on the keywords that are found in the recognition text.
 - This identification technique requires recognition text and full page recognition enabled.
- Keyword text file
 - A list of words or phrases that are separated by new lines, in a file.
 - Word matching is case-sensitive.
 - The file must have a “.key” extension for the system to recognize it as a keyword file.
 - The file is used for matching.
- Examples of keywords:
 - Donation, invoice, receipt, ticket

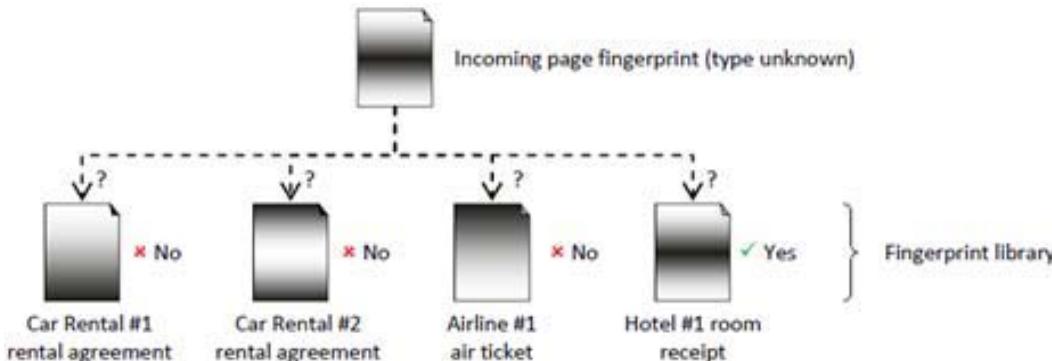
Figure 1-27. Page identification – Keywords

How do keywords work in page identification?

- The search first looks for the first word or phrase in the keyword file.
- Common substitutions are applied to search criteria to improve results.
- Starting from the location of the last find, if the word is found, the search stops.
- If no match is found, the next line from the keyword file is read and again the search starts from the result of a previous find.
- This process continues until a match is found or all of the lines in the keyword file are read.
- The location of the found word or phrase that matches an entry in the keyword file is remembered to be used by subsequent actions.

Page identification – Fingerprints

- Fingerprinting
 - Identifies the current page based on fingerprint matching.
 - Datacap generates a fingerprint that describes each incoming page.
- A fingerprint consists of an image file (.tif) and a recognition file (.cco).
- The CCO contains the location of all words and lines.
- Example: Compare dark and light zones to identify an incoming page.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-28. Page identification – Fingerprints

How does the Fingerprinting technique work?

- The fingerprint can include information about the relative densities of different regions of the page or the location of text on the page.
- Datacap compares an incoming page fingerprint to existing ones.
- If it matches an existing fingerprint, it is safe to assume that the incoming page is of the same class as the existing one.
 - The offset required to give the best match is also captured.
- If a match does not occur, it creates a fingerprint.
- Technique is adapted for structured and semi-structured documents with a fairly constant layout.

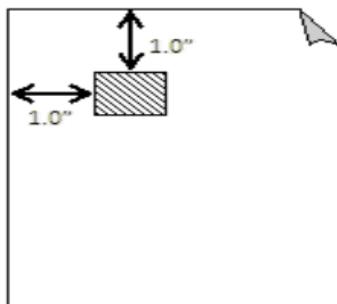
Example of page identification

In the example that is shown on the diagram, the incoming page matches the Hotel #1 room receipt.

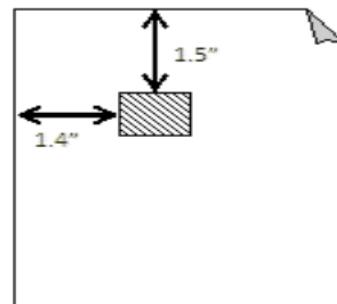
Datacap assigns it the type “Room_Receipt”, and records the ID of the matching fingerprint in the runtime batch hierarchy.

Page identification – Pattern Recognition

- Pattern recognition
 - Identifies the current page based on geometric patterns such as:
 - Vendor logos.
 - Page registration marks.
 - Text-based patterns.
 - You can use Datacap pattern matching to identify pages and adjust misaligned or distorted images.



Fingerprint



Scanned page

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-29. Page identification – Pattern Recognition

Text analytics – (Content Classification) You must enable full page recognition.

Other Page identification methods

- Manual ID
 - Manual page ID selection by the operator.
- Separator sheets
 - Identifiable separator sheets are physically placed in the batch so that they can be used to identify the following page.

Page identification – Text analytics

- Text analytics (Content Classification)
 - Identifies the current page by using the IBM Content Classification Knowledge Base.
 - Analyzes text to try to find matches.
- Enable Datacap applications to use the IBM Content Classification Knowledge Base for fingerprint matching.

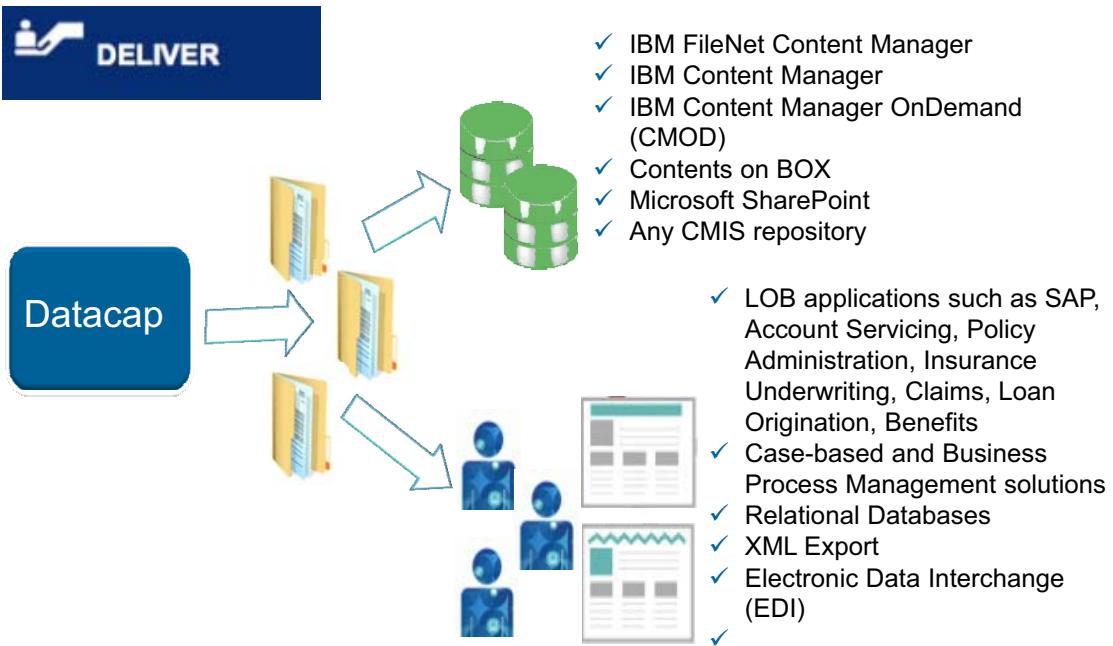
Figure 1-30. Page identification – Text analytics

Text analytics – (Content Classification) You must enable full page recognition.

Other Page identification methods

- Manual ID
 - Manual page ID selection by the operator.
- Separator sheets
 - Identifiable separator sheets are physically placed in the batch so that they can be used to identify the following page.

Deliver: Captured Documents and Data



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-31. Deliver: Captured Documents and Data

Medical Claims Datacap application uses Electronic Data Interchange (EDI)

Datacap Capture process

- Organize the flow of tasks in the capture process from scan to export and exceptions in a workflow.
 - Split batches to group documents for exception or prioritized processing.
 - Control access to the system and tasks that use its security features.
 - Monitor progress of capture operations and fix problems in real time.
 - Report on capture operations and provide statistics on how well the system is doing.
 - Run unattended tasks in the background that uses the Rulerunner service
 - Provide a unified execution environment for background processing
 - Run multiple process threads to increase throughput.
 - Support flexible deployment scenarios.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-32. Datacap Capture process

Support flexible deployment scenarios

Example: Central mail room-type of operations and distributed imaging over the web, or at regional offices.

Rules and rulesets

- Rules
 - Rules are assigned to process specific objects (Example: Analyze and identify pages).
 - Entity that is tied to a Datacap Object (DCO Objects are: Batch, Document, Page, or Field).
 - An ordered set of functions that process an object.
 - Reusable and extensible.

- Ruleset
 - A ruleset consists of one or more rules.

Figure 1-33. Rules and rulesets

Data verification

- Define rules to recognize and validate image content.
 - Based on the validation results, you decide to route the data to an operator for manual verification.
- In the verification step, the operator corrects and validates the fields.
- The following user interfaces allow manual verification:
 - Datacap Navigator.
 - Datacap Desktop.
 - FastDoc.
- You can customize the verification (panels) interface.

Figure 1-34. Data verification

Review questions (1)

1. The Scanners and Multi-Functional Devices input channels support which of the following file types?
 - Select more than one option:
 - A. TIFF
 - B. JPEG
 - C. TXT
 - D. HTML
 - E. PDF
 - F. DOCX
 - G. ZIP

2. True or False.

Both one- and two-dimensional barcodes are used for page recognition.

Figure 1-35. Review questions (1)

Review answers (1)

1. The Scanners and Multi-Functional Devices input channels support which of the following file types?
 - Select more than one option:
 - A. TIFF
 - B. JPEG
 - C. TXT
 - D. HTML
 - E. PDF
 - F. DOCX
 - G. ZIP

The answers are : A, B, and E

2. True or False.

Both one- and two-dimensional barcodes are used for page recognition.

The Answer is: True

Figure 1-36. Review answers (1)

Review questions (2)

3. True or False.

Datacap captured documents and data can be exported to content repositories or can be used in applications.

4. Which of the following items are Datacap page identification methods?

Select more than one option:

- A. Keyword
- B. Batch Process
- C. Pattern recognition
- D. Fingerprint
- E. Document export
- F. Input channel

Figure 1-37. Review questions (2)

Review answers (2)

3. True or False.

Datacap captured documents and data can be exported to content repositories or can be used in applications.

The answer is: True.

4. Which of the following items are Datacap page identification methods?

Select more than one option:

- A. Keyword
- B. Batch Process
- C. Pattern recognition
- D. Fingerprint
- E. Document export
- F. Input channel

The Answers are: A, C, and D

Figure 1-38. Review answers (2)

Lesson 1.3. Role-based Datacap clients

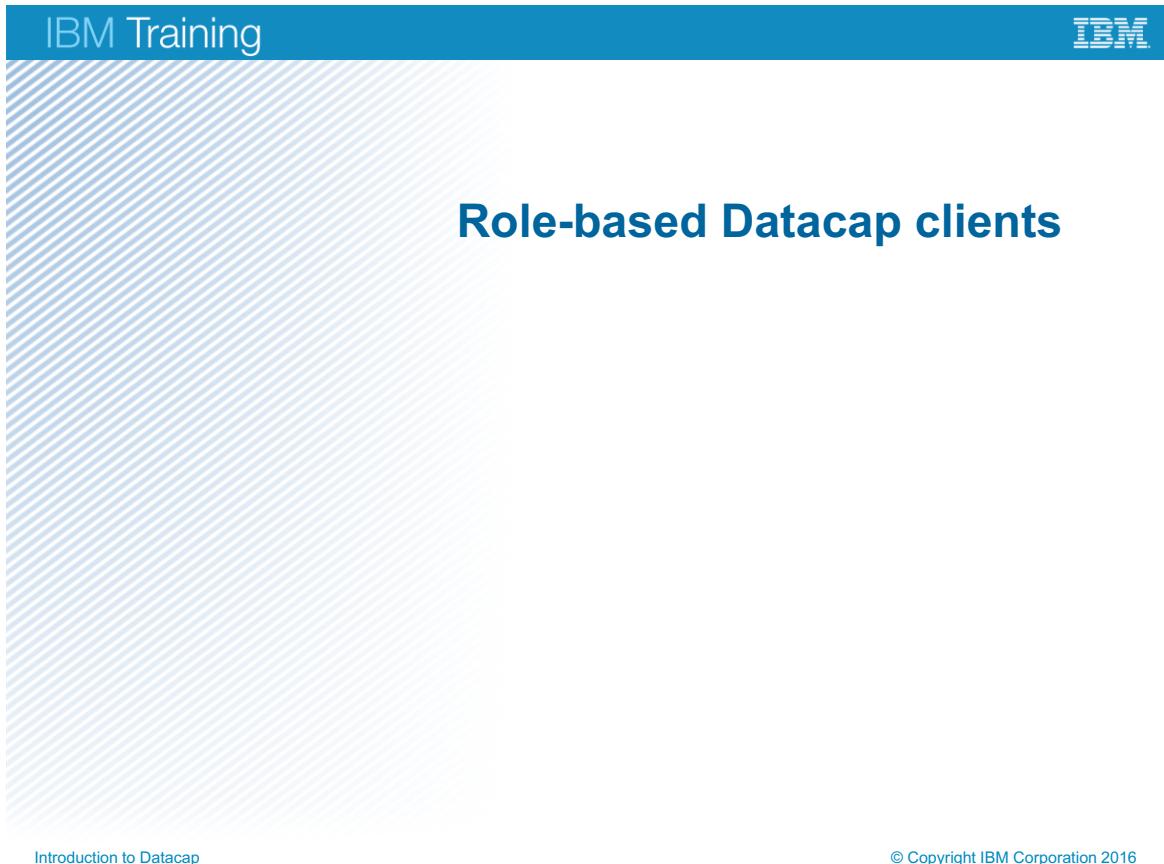


Figure 1-39. Role-based Datacap clients

Lessons

- Datacap overview
- Datacap process
-  Role-based Datacap clients
 - Architecture configurations
 - Architecture components
 - Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

Figure 1-40. Lessons

Why is this lesson important to you?

- This lesson provides an overview of different Datacap clients for different business roles.

Figure 1-41. Why is this lesson important to you?

Role-based Datacap tools

- Business users
 - Datacap Navigator – Web client
 - Datacap Mobile
 - Datacap Desktop – Windows based client
 - Datacap FastDoc – Windows based client
- Business Analysts
 - FastDoc (for configuring applications)
- Developers
 - Datacap Studio (for configuring and developing applications)
- Administrators
 - Datacap Navigator (new in Datacap 9.0)
 - Server management tools

Figure 1-42. Role-based Datacap tools

More details and hands on labs for each tool in the list are provided in the following lessons.

Datacap Navigator

- Datacap client for scanning, uploading, manual classification, and administration of the workflow and users.

Datacap tmweb client

- Datacap client for scanning, uploading, manual classification, and administration of the workflow and users.
- tmweb is being phased out in favor of Datacap Navigator web client.

Datacap Desktop

- Datacap client for scanning, manual classification, user data correction, and manual background processing for testing background tasks.

Datacap FastDoc

- Uses application templates and prebuilt rulesets to quickly configure functional applications.
- Datacap client for scanning, manual classification, user data correction, and manual background processing for testing background tasks.

Datacap Studio

- A more robust development environment to customize FastDoc built applications or build custom applications from scratch.

IBM Training

The screenshot displays three main modules of the Datacap Navigator - Web client:

- Scan:** Shows a preview of a scanned document titled "Car Rental #1" and its "Watch Structure".
- Monitor:** Displays a "Job Monitor" table with columns: Queue ID, Rank, Job, Task, Status, and Job Start. It lists various jobs with details like "NavigateVerify pending" or "NavigateVerify failed".
- Verify:** Shows an "Image Viewer" with a "Field Details" panel containing flight information: Outbound From (New York/Newark (EWR)), Outbound To (San Francisco (SFO)), Return From (San Francisco (SFO)), and Return To (New York/Newark (EWR)).

At the bottom right, the text "© Copyright IBM Corporation 2016" is visible.

Figure 1-43. Datacap Navigator – Web client

IBM Training

Datacap Mobile Capture

- Capture images
 - From mobile device directly into Datacap application.
 - Auto or manual mode.
- Rich on-device feature set
 - Automatic edge detection, deskew, and snap when quality thresholds met.
 - Crop, contrast/brightness, rotate, and reorder
- On-device classification
- Indexing (data entry)
 - On-device zonal OCR
 - On-device barcode recognition
- Submission of the document to the Datacap server.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-44. Datacap Mobile Capture

IBM Datacap Mobile app captures images into Datacap application at the point of origination. The app processes, and uploads documents to a Datacap server.

Capture Images

- Capturing in Auto mode:
 - Is the default configuration for Datacap Mobile.
 - Reduces the learning curve for the user, and to ensure the user gets the best possible results.
- Capturing in Manual mode
 - It is not possible to use automatic mode for all occasions.
 - Manual mode is suitable when there is no notion of a “page” with edges or minimum levels of quality for OCR.

Example: Photographic evidence of property damage.

 - Some content might need to be submitted as evidence along with a main document.

Example: Receipts.

Classification, or identification

It is the process of determining what pages we are working with and how they fit together into a document.

Indexing

- It is the final step in preparing the document for submission to the Datacap server is to extract data from the images into designated fields.
- Datacap Mobile supports three types of indexing:
 - Manual
 - On-device OCR
 - On-device Barcode recognition
- Indexing can be done at all levels of document hierarchy. Only page level is exposed in the default app.
- Although indexing can be done on the server-side, Datacap Mobile allows it at the point of capture. The user can ensure that the indexing is done accurately and comprehensively to minimize any errors later in the business process.

Submission

Submit the document to the Datacap server for further processing.

IBM Datacap Mobile app

- Compatible platforms:
 - Apple iOS
 - Google Android
- Compatible software versions:
 - iOS 8.0.0 or later (iPhone or iPad)
 - Android 4.4 or later (Android smartphone or tablet devices)
 - A licensed version of IBM Datacap at version 9.0.0.1 or later
 - Mobile app is available on Apple App Store, iTunes and Google Play
- App developers can customize and extend the mobile app with the Datacap Mobile SDK APIs and documentation.



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-45. IBM Datacap Mobile app

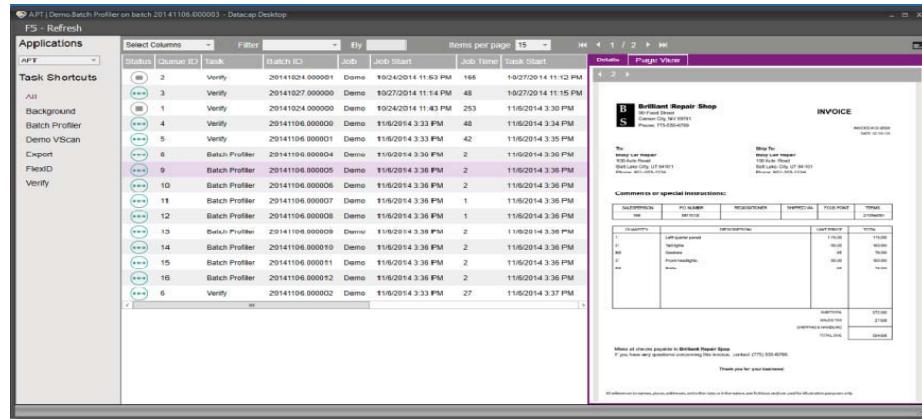
Help Path

- Datacap 9.0.1>Installing and configuring in a client/server environment>Datacap installation and configuration in a client/server environment>Installation instructions for Datacap server>Datacap Mobile app configuration

http://www-01.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.dc.install.doc/dcain447.htm?lang=en

Datacap Desktop – Windows based client

- Datacap Desktop
 - Scan (with high-speed scanners)
 - Customizable Verification Panel
 - Monitor tasks
 - Custom batch fields support
 - Full batch content display



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-46. Datacap Desktop – Windows based client

Custom Verification Panel - Datacap Desktop provides an interface that you can customize for verifying and correcting information from scanned documents, or for manually entering information.

IBM Training



Datacap FastDoc – Windows based client

- Rapid applications development in a stand-alone environment.
 - Configure applications to scan, index, and manually run background tasks on documents quickly without using Datacap Studio.
- As a client (to scan and verify) to Datacap.

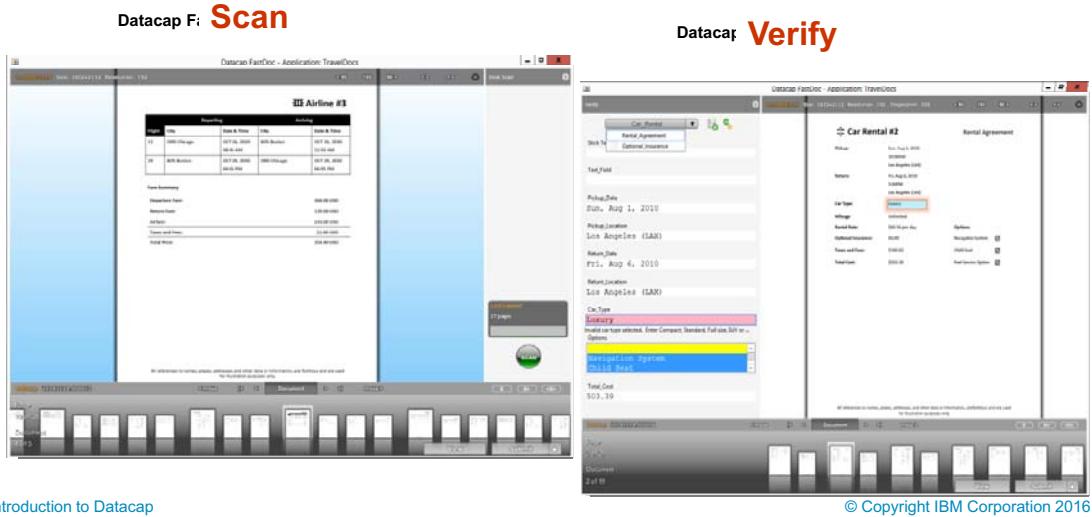


Figure 1-47. Datacap FastDoc – Windows based client

Rapid application development in FastDoc

Configure new Datacap applications quickly without using Datacap Studio.

You can configure applications to scan, index, and manually run background tasks on documents.

FastDoc as a Datacap client

FastDoc runs as a client that scans, auto indexes, and uploads batches of documents to Datacap Server.

Datacap Applications

- IBM Datacap installation includes a sample application:
 - [TravelDocs](#)
- The following applications are available as separately licensed:
 - [Datacap Accounts Payable](#)
 - [Datacap Medical Claims](#)
- Examples of Industry-specific applications that are created with Datacap:
 - [Loan Applications](#)
 - [Census forms](#)

Figure 1-48. Datacap Applications

Datacap Accounts Payable

- Capture and verify invoice data without manual data entry
- Captures all line items, even on multi-page invoices
- Learns new invoice types continuously
- Validation rules on dates, math, lookups, data types
- Aids three-way match with purchase order line item reconciliation
- SAP integration solution that is provided by Business Partners

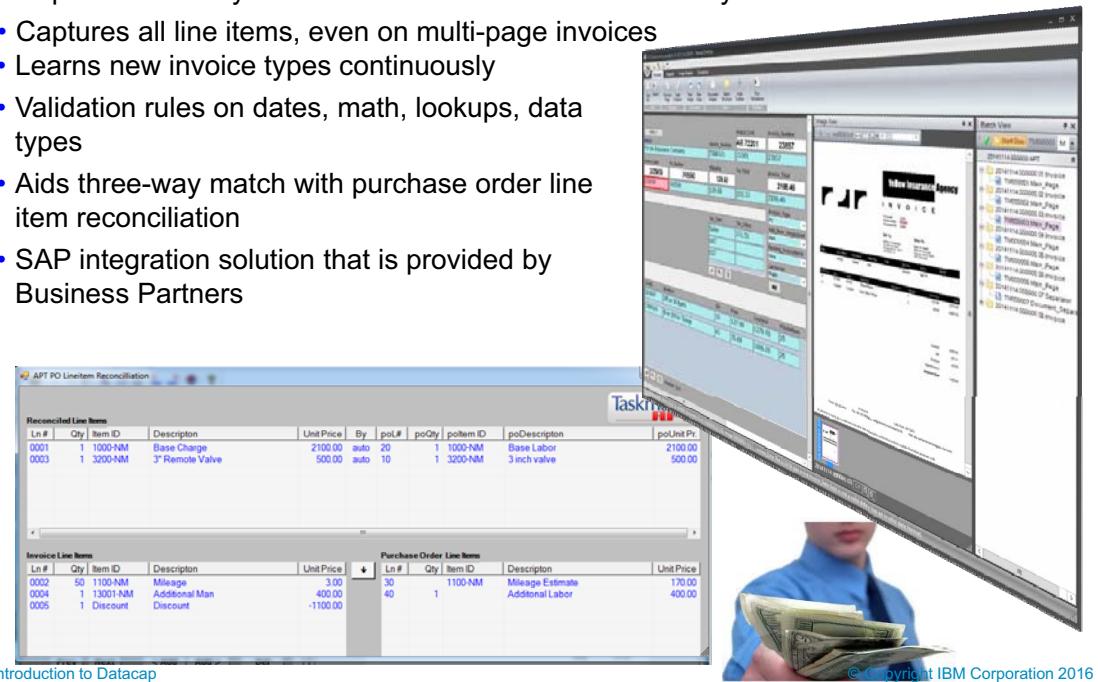


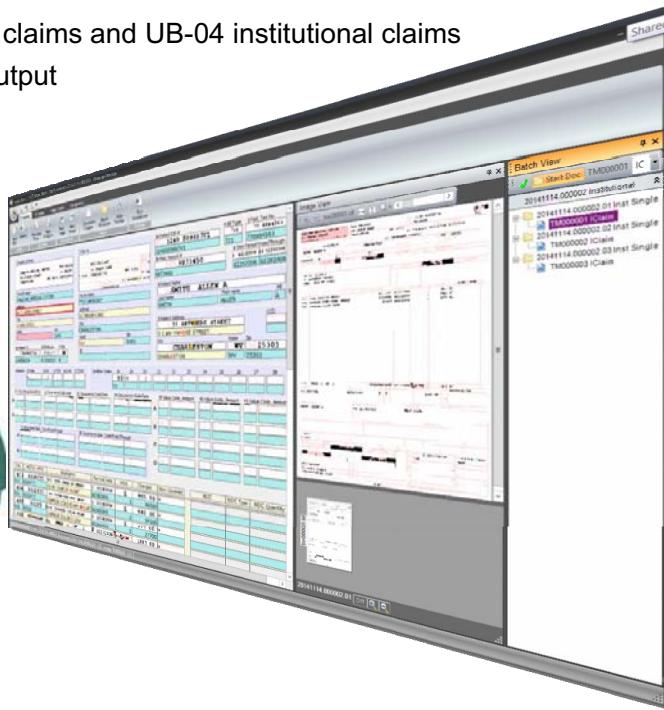
Figure 1-49. Datacap Accounts Payable

IBM Training



Datacap Medical Claims

- Capture CMS 1500 medical claims and UB-04 institutional claims
- HIPAA-compliant 837 EDI output
- Capture all fields
- Plus attachments
- Support for black claims
- Healthcare Validations



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-50. Datacap Medical Claims

Role-based Datacap clients

English, French, Spanish,
German, Dutch, Italian,
Portuguese, Swedish,
Russian, Hungarian, Polish,
Romanian, Czech, Slovak,
Turkish, Greek, Arabic,
Hebrew, Chinese, Japanese



52

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-51. Role-based Datacap clients

For more information on “IBM Datacap 9.0.1 Language Support”, please see:

<http://www-01.ibm.com/support/docview.wss?uid=swg27044111>.

Review questions

1. Which one of the following items is not a Datacap client for business users to process a batch?
 - A. Datacap Navigator
 - B. Datacap Mobile
 - C. Datacap Desktop
 - D. Datacap FastDoc
 - E. Datacap Studio
2. True or False.
Datacap FastDoc can be used to rapidly configure the Datacap applications and as a client to scan and verify the documents.

Figure 1-52. Review questions

Review answers

1. Which one of the following items is not a Datacap client for business users to process a batch?
 - A. Datacap Navigator
 - B. Datacap Mobile
 - C. Datacap Desktop
 - D. Datacap FastDoc
 - E. Datacap Studio

The Answer is: E

2. True or False.

Datacap FastDoc can be used to rapidly configure the Datacap applications and as a client to scan and verify the documents.

The Answer is: True

Figure 1-53. Review answers

Lesson 1.4. Architecture configurations

Architecture configurations

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-54. Architecture configurations

Lessons

- Datacap overview
- Datacap process
- Role-based Datacap clients
-  Architecture configurations
 - Architecture components
 - Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

Figure 1-55. Lessons

Why is this lesson important to you?

- This lesson provides an overview of the configuration variations for a Datacap Capture system.

Figure 1-56. Why is this lesson important to you?

Architecture configurations

- Single system configuration – Components are all installed on the same system.
 - Used for providing product demonstrations, in a proof-of-concept environment, or during initial product evaluation.
- Client/server configuration – Components are installed on dedicated systems.
 - Supports up to hundreds of simultaneous users, and uses centralized application management and shared databases.
- Hybrid Solution – Components are installed on shared systems.
Examples:
 - Datacap web client (tmweb), Report Viewer, and Datacap Web Services (wTM) sharing one system.
 - Rulerunner service and Fingerprint service sharing one system.

Figure 1-57. Architecture configurations

Datacap Capture can be deployed in various configurations.

Single system architecture

At one end of the spectrum is the single-system configuration, where Datacap Capture components are installed on the same system. This configuration is typically used for providing product demonstrations, in a proof-of-concept environment, or during initial product evaluation.

Single-system installations can also be used for development systems and for training systems like the one used for this class.

Using a single system architecture, all of the Datacap Capture components can be on the same system. You can install Datacap with various content management systems, described in the introduction lesson, to provide direct access to data repository.

Client/server configuration

At the other end of the spectrum, is the client/server configuration, where the various Datacap Capture components are installed on dedicated systems. (web servers and database servers). Client/server configurations can support up to hundreds of simultaneous users and uses centralized application management and shared databases.

Hybrid Solutions

And spanning the center of the spectrum are various hybrid configurations in which two or more Datacap Capture components are installed on the same system. You might, for example, run Datacap Web (tmWeb), Report Viewer, Datacap web services on the same web server. You might also install and run the Datacap Rulerunner Service and the Fingerprint Service on another server. In any production environment, it is considered good practice Datacap Rulerunner service on a dedicated server because it is very CPU intensive.

Architecture configurations

- Server/Client and hybrid configurations that are classified geographically.
 - Centralized Deployment.
 - All Datacap services are done at a central site.
 - Distributed Deployment.
 - Services are distributed to two or more sites.

Examples

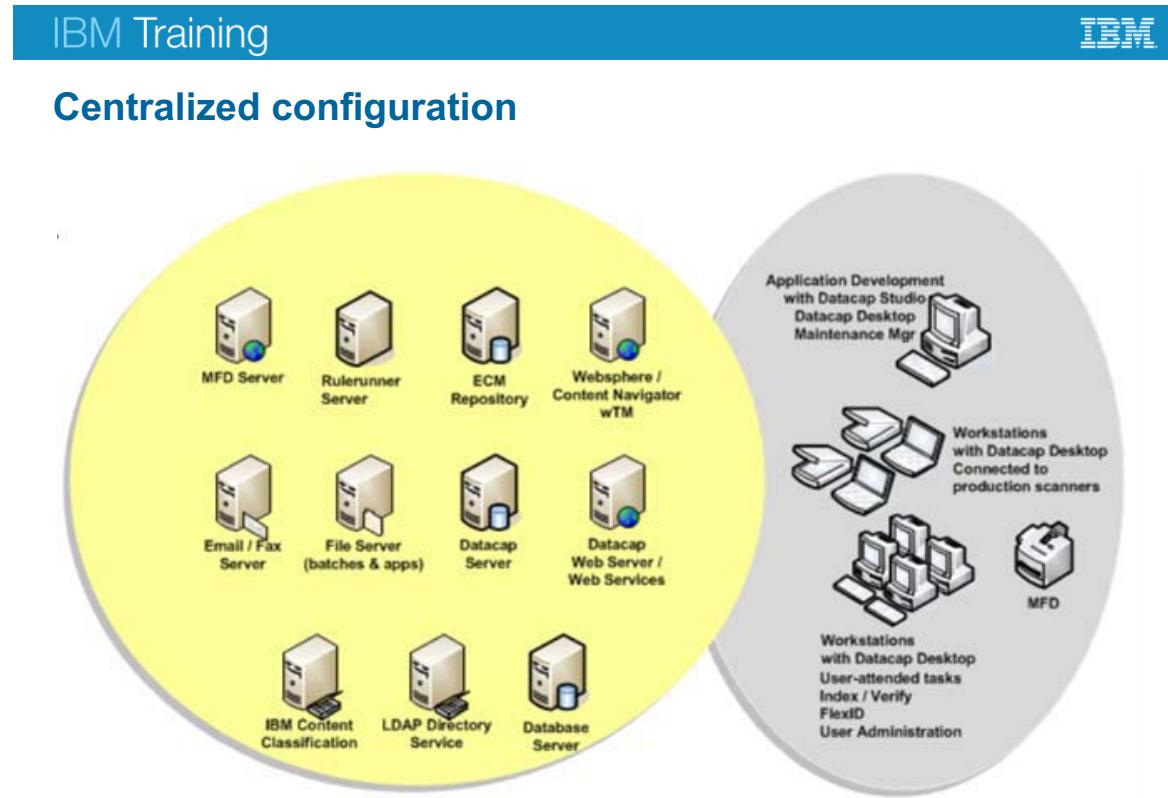
- Centralized capture
- Distributed Scanning with centralized Indexing
- Distributed capture with local indexing and central control.

Figure 1-58. Architecture configurations

The client/server and hybrid configurations can be further categorized geographically into more configurations:

Some configurations might include these examples:

- Centralized capture
 - All functions that are done at a central location, that is a one geographic location.
- Distributed Scanning with centralized Indexing
 - The scanning function is done at a remote site and that batches are uploaded to the central location for indexing and further processing.
- Distributed capture with local indexing and central control.
 - The scanning and indexing functions are done at remote locations and that is then uploaded to a central site for exporting to a repository, archiving, or further processing.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-59. Centralized configuration

Centralized Configuration

- This diagram represents a centralized configuration where all functions are done at a central site.
- The systems in the yellow area are in a server room.
- The workstations, production scanners, and MFD (Multi-Function Device) scanners are at user accessible stations or offices and cubicles.



Distributed configuration

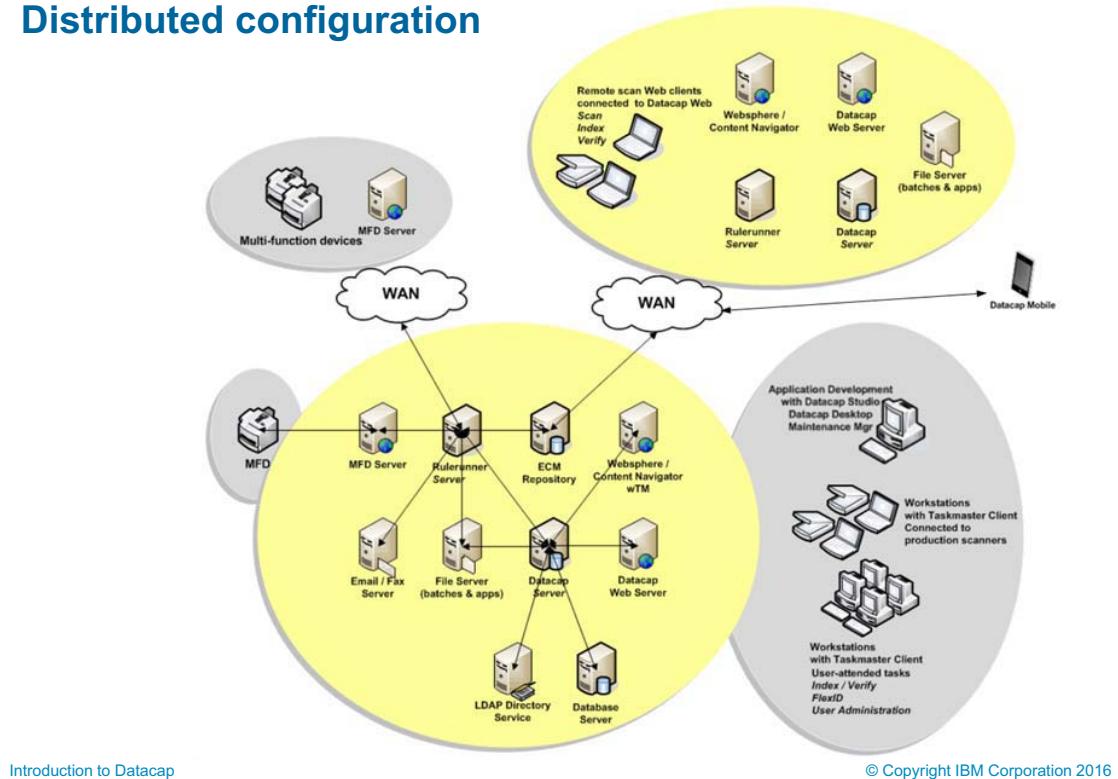
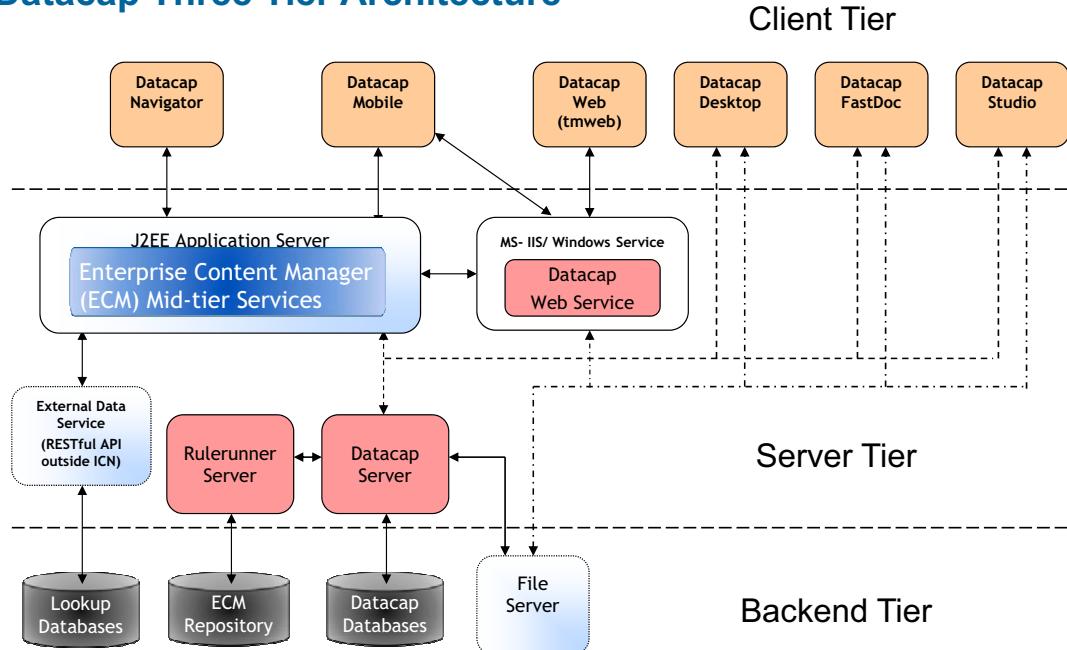


Figure 1-60. Distributed configuration

Distributed Configuration

- This diagram represents a distributed configuration where all functions can be done at the central site.
- A remote satellite site provides services for scanning and indexing but database, LDAP, and Enterprise Content Manager Repository services are provided by the main site.
- A third remote site that is scanning only station with multi-function devices and a multi-function server would upload the scanned batches to the main site for indexing and other processing.

Datacap Three Tier Architecture



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-61. Datacap Three Tier Architecture

Datacap Three Tier Architecture

The Datacap components are designed to operate in a three tier configuration.

Client Tier

The client tier consists of all of the components that are hosted on workstations or mobile devices where humans interface to the Datacap system.

Server Tier

The server tier is where all the Datacap services are provided. This interface is used between the users and the backend storage devices. The main service components are:

- Datacap Service – Authentication, batch task-queuing, database access, and file sharing.
- Rulerunner Service – Background task processing.
- Web Services – For web and mobile client connection. RESTful API services for external access to Datacap services and also used by IBM Content Navigator to provide Datacap Navigator and Mobile connectivity.

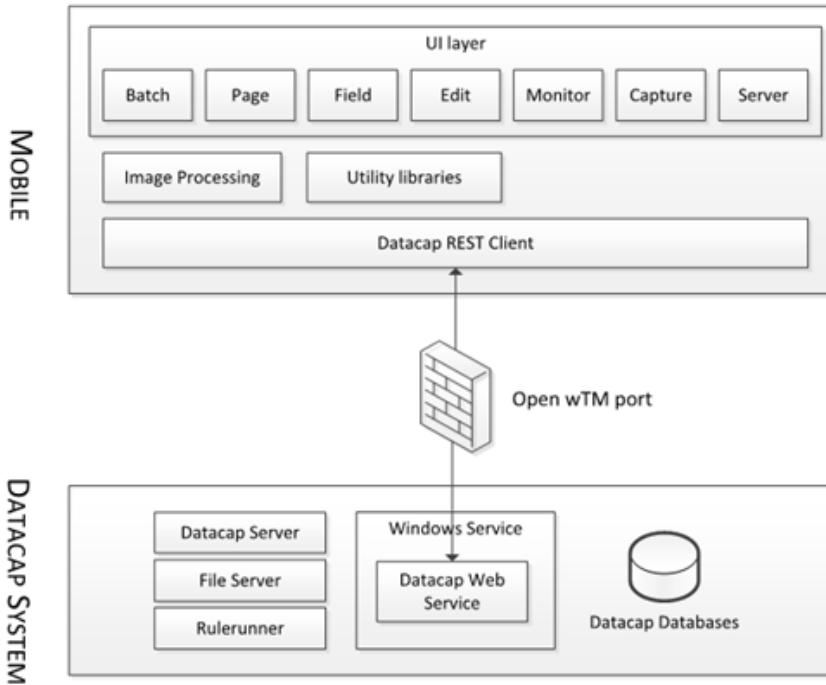
Backend Tier

The backend tier is where storage servers are configured.

- Datacap databases are:
 - Engine for task queueing batch monitoring,
 - Admin for Datacap authentication,
 - Fingerprint for fingerprint storage and management.
 - Lookup for validating or populating metadata fields.
- File Server
 - Storage of Datacap applications and other control code, libraries, and tools.
 - Storage of Fingerprint images and data.
 - Temporary storage of batch task information, images, and logs.
- ECM Repository
 - Permanent Storage of images and searchable metadata for Datacap export data.
- External databases and services.
 - Example: lookup databases



Mobile configuration



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-62. Mobile configuration

Datacap mobile architecture with remote access

This illustration shows the components of Datacap components that are hosted in a mobile device and how the mobile device connects to the Datacap services.

Mobile

- IBM Content Navigator Mobile application with Datacap Document capture.
- Connects to the Datacap System through a firewall protected open Datacap Web Services (wTM) port.

Datacap Systems

- Datacap Web Services (IIS)
- Datacap Server
- File Server (applications and batches)
- Rulerunner Server
- Datacap Databases

Review questions (1)

1. True or False

Many production Datacap installations can be configured with a Single system configuration.

2. True or False

The most efficient and cost effective Datacap installations are client/server configurations where all software components are installed on dedicated servers.

3. True or False

Most production Datacap systems have some Datacap components that are installed on dedicated systems and some components on shared systems.

Figure 1-63. Review questions (1)

Review answers (1)

1. Many production Datacap installations can be configured with a Single system configuration.

The answer is: False

2. The most efficient and cost effective Datacap installations are client/server configurations where all Datacap software components are installed on dedicated servers.

The answer is: False

3. Most production Datacap systems have some Datacap components that are installed on dedicated systems and some components on shared systems.

The answer is: True

Figure 1-64. Review answers (1)

Review questions (2)

4. Consider the terms “Centralized Deployment” and “Distributed deployment”. Which of the following statements is correct.
- A. Centralized and Distributed refer to how Datacap components are deployed across servers in a Datacap configuration.
 - B. Centralized and Distributed refer to geographic location of Datacap services and tasks.

Figure 1-65. Review questions (2)

Review answers (2)

4. Consider the terms “Centralized Deployment” and “Distributed deployment”. Which of the following statements is correct.
- A. Centralized and Distributed refer to how Datacap components are deployed across servers in a Datacap configuration.
 - B. Centralized and Distributed refer to geographic location of Datacap services and tasks.

The answer is: B

Figure 1-66. Review answers (2)

Review questions (3)

5. Consider the scenario where all of the Datacap services are provided by servers that are in a single-server room. Scanning and verification tasks are done from a workstation or scanning stations throughout multiple buildings at the same physical address and all connected to the same LAN. What is the classification for this scenario?
 - A. Centralized
 - B. Decentralized
6. Consider the scenario where all of the Datacap servers are at one physical location. Scanning and verification tasks are done from the location and from remote locations that are connected over the internet. What is the classification for this scenario?
 - A. Centralized
 - B. Decentralized

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-67. Review questions (3)

Review answers (3)

5. Consider the scenario where all of the Datacap services are provided by servers that are in a single-server room. Scanning and verification tasks are done from a workstation or scanning stations throughout multiple buildings at the same physical address and all connected to the same LAN. What is the classification for this scenario?
- A. Centralized or B. Decentralized
- The answer is: A
6. Consider the scenario where all of the Datacap servers are at one physical location. Scanning and verification tasks are done from the location and from remote locations that are connected over the internet. What is the classification for this scenario?
- A. Centralized or B. Decentralized
- The answer is: B

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-68. Review answers (3)

Lesson 1.5. Architecture components



Figure 1-69. Architecture components

Lessons

- Datacap overview
 - Datacap process
 - Role-based Datacap clients
 - Architecture configurations
-  **Architecture components**
- Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-70. Lessons

Why is this lesson important to you?

- This lesson provides an overview of the Datacap system components.

Figure 1-71. Why is this lesson important to you?

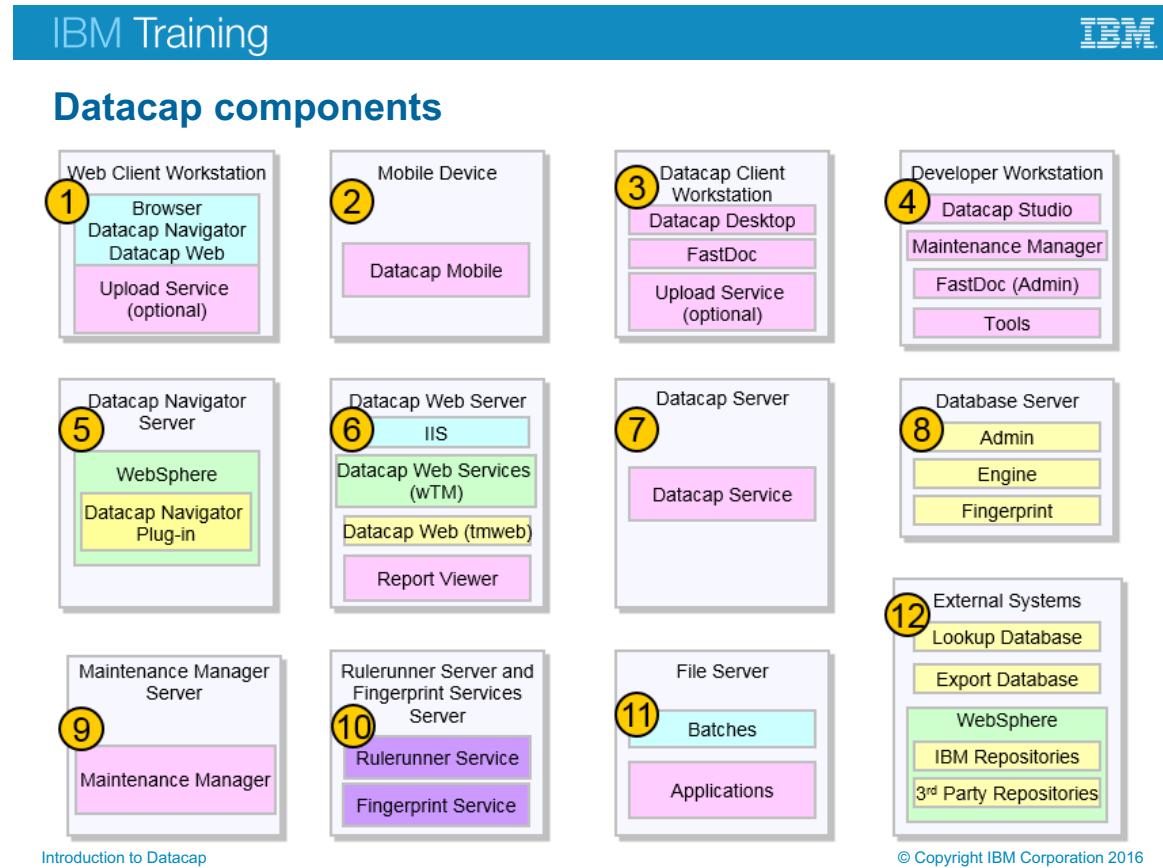


Figure 1-72. Datacap components

Datacap Components

This diagram represents one possible mapping of Datacap software components to physical server or workstation computers. There can be many other valid combinations.

1. Web Client Workstation with browser and Upload Service (optional). Datacap Navigator and Datacap Web (tmweb) clients that are run in the browser. This configuration is used for a remote user station or a remote scanning station.
2. Datacap Mobile is a iOS or Android-based Mobile device that is configured with apps that provide connectivity the Datacap system.
3. Datacap Client Workstations with Datacap Desktop, FastDoc, and Upload Service (optional). This configuration is for a local user or a local scanning station.
4. Developer Workstation with Datacap Studio, Maintenance Manager, FastDoc (Admin), and Tools. This configuration is for a business analyst, application developer, or system administrator.
5. IBM Navigator Server that is configured with the Datacap Navigator plug-in.

6. Datacap Web Server (IIS)

Note: Datacap Web Services wTM is shown here installed under IIS. It can also be installed directly as a windows service.

7. Datacap Server

8. Database Server (Admin, Engine, and Fingerprint)

9. Maintenance Manager Workstation (Maintenance Manager)

10. Rulerunner Server and Fingerprint Services Server

11. File server (Batches & Applications)

12. External Systems:

- a. Lookup Databases (Example: vendor, customer, and purchase order)
- b. Export database
- c. IBM Repositories
- d. Third Party Repositories

Components - Datacap Navigator and Datacap Web Client

- Used for remote business users for scanning, verification, administrative tasks, and application configuration.
- Browser-based client runs on available browser.
 - Datacap Web Client (tmweb).
 - Datacap Navigator.
- Communicates with Datacap Web servers for all services.
- Session cookies, require session affinity.
- ActiveX controls are used for scanning, image viewing, snippets, data entry fields, upload.
- Scan – supports TWAIN drivers.
- Multiple user interface layouts and functions support varying use cases.

Figure 1-73. Components - Datacap Navigator and Datacap Web Client

Datacap Navigator and Datacap Web Client

- These web clients provide functions similar to the Datacap client but does not require more software to be installed on the workstation.
- Multiple user interface layouts and functions support varying use cases.
 - Scan
 - Upload
 - Classification
 - Indexing
 - Verification
 - Multi-pass
 - Double blind
 - Push and pull
 - Administration

Datacap Web Client

- When you verify a batch by using the web client, verification rules are run on the web server. You can also configure an application workflow and run administrative tasks such as setting up Datacap groups and users.

Datacap Navigator

- Authentication is done through calls to Datacap Web Services.
- Is based on IBM Content Navigator technology and is installed, configured, and administered with IBM Content Navigator tools.
- Datacap Navigator communicates with the Datacap Server by using the Datacap Web Services APIs.
- For lookup and verification, Datacap Navigator uses the IBM Content Navigator External Data Services infrastructure.

Components – Datacap Client Workstation

- Used by business users to run tasks such as scanning and verification with one of the installed clients.
- When verifying a batch, verification rules are processed on the Datacap Client Workstation.
- Uses Taskmaster Server for queuing and security
- Accesses the file server to process images
- Accesses the fingerprint database and lookup databases through ODBC
- Scan supports ISIS drivers
- User windows based clients are Datacap Desktop, FastDoc.
- Browser hosted thin clients are Datacap web client (tmWeb), and Datacap Navigator.

Figure 1-74. Components – Datacap Client Workstation

Datacap client

The Datacap client component is a set of programs that provide user access to Datacap applications. The Datacap Desktop and FastDoc user interfaces are Datacap software components that run on Windows to run the user tasks, such as scanning and verification. When a user verifies a batch, the verification rules are run on the Datacap client workstation.

The Datacap Client Workstation component includes:

- Datacap Desktop for task processing and batch monitoring.
- Datacap FastDoc for offline and online task processing.
- Optional browser hosted Datacap web client (tmWeb), and Datacap Navigator.
- Optional Upload Service for automatically uploading scanned batches.

Components – Developer Workstation

- Used by application builders, developers, and administrators.
- Installed tools and applications:
 - Datacap Studio.
 - Datacap FastDoc.
 - Datacap Maintenance Manager
 - Application Manager
 - Datacap Report Viewer
 - Other Tools

Figure 1-75. Components – Developer Workstation

Datacap Studio

- Is the Datacap application development environment.
 - Configures rules, actions, fingerprints, document hierarchies, fields

Datacap FastDoc

- You can create applications on FastDoc to scan, index, and manually run background tasks on documents in a stand-alone environment or as a client to Datacap.

Datacap Maintenance Manager

- Provides application monitoring and notification capabilities.
- Monitor batches, send alerts, and do automate housekeeping tasks.
- Used to delete completed batch records and folders to free disk space.

Application Manager

- To manage multi-system distributed environments through a centralized set of key Datacap configuration settings that are stored in shared files.

Report Viewer

- Report Viewer is the reporting tool for real-time reports of Datacap activity. Report Viewer gets usage statistics and other data from the Engine database.

Other tools

- Application Wizard
- Database creation and copy utility.

Components – Datacap Navigator Server

- Datacap Navigator Server hosts web applications with IBM WebSphere Application Server.
- Datacap Navigator is based on IBM Content Navigator technology.
- Detailed configuration and functions information was already provided in the earlier Datacap navigator lesson.

Figure 1-76. Components – Datacap Navigator Server

Components – Datacap Web Server

- Datacap Web Server system hosts Datacap web applications with Microsoft IIS application server.
- Datacap web service components are:
 - Datacap Web Server (tmweb server).
 - Report Viewer Server (RV2 server).
 - Datacap Web Services (wTM - RESTful web service interface).
- Communications from IIS to back-end services is through the Datacap Server.

Figure 1-77. Components – Datacap Web Server

Datacap Web Server

- When scanning and verifying a batch with the web clients, the rules are run on the Datacap Web server.
- For tasks that do not require operator intervention, the application is configured to use the Rulerunner Service for background tasks.
- Validation rules are run on the Datacap Web server.
- The Datacap web service components are:
 - Datacap Web Clients that are serviced by Datacap Web Server are:
 - Datacap Web Client (tmweb).
 - Can do Datacap batch scanning and verifying, monitoring, administration, and configuration tasks.
 - Report Viewer Server (RV2 server).
 - Is the reporting tool that shows real-time reports of Datacap activity.
 - Gets usage statistics and other data from the Engine database.

- Provides a set of standard reports and the ability to customize existing reports and create new reports.
- The following standard reports for monitoring batch status, station activity, and problem batches are included with the Report Viewer software component.
- Can be collocated with Datacap Web server.
- RESTful web service interface for custom applications and services.

Components – Datacap Web Services

- Datacap Web Services are also called wTM.
- Windows service or Microsoft IIS-based web service.
- For interaction with Datacap through a simple REST API.
 - Create, process, and release a batch.
 - Get information about a batch.
 - Process documents outside of a Datacap batch.
- Datacap Web Services support HTTP and HTTPS protocols.

Figure 1-78. Components – Datacap Web Services

Help path

- Datacap 9.0.1>Reference>Datacap Web Services REST API methods
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.dc.install.doc/dcdws031.htm

Components – Datacap Server

- Controls authentication, database access, and batch-queuing.
- Accesses the administration and engine databases.
 - All other components must access the databases through Datacap Server.
- Accesses the file system for thin clients.
 - Navigator and web services clients access application and document files through the Datacap server.
- Datacap Servers can be deployed in an active - active farm.
- Runs as a Windows service.

Figure 1-79. Components – Datacap Server

Help path

- Datacap 9.0.0 > Overview > Datacap software components

Datacap Server

- Datacap Server service runs as a Windows service and controls authentication, database access, and controls batch-queuing.
- The main application settings file (datacap.xml) is on the Datacap Server.
- Only the Datacap Server can access the administration and engine databases.
- All other components must access these databases through Datacap Server.
- Provides access to the application and document files for the Datacap Navigator and Web services clients.
- Load-balancing and farming Datacap servers.
- For high volume installations, servers can be configured in server farms where all servers in the farm are active (Active-active mode).
- You can use network load balancers to manage client requests across servers in a Datacap system.

- Load-balancing is a method for scaling a system horizontally by distributing the work across many computer nodes in a "farm".
- It also provides high availability by redirecting clients to a working node in case of failure.

Components – Datacap Database Server

- Supported Relational databases.
 - Oracle, Microsoft SQL Server, DB2, Microsoft Access
- Datacap databases.
 - Admin: Stores information about workflows, users, and groups.
 - Engine: Stores batch information and statistics.
 - Fingerprint: Stores page template data.
Note: Optionally, Admin, Engine, and Fingerprint can be tables in a single database.
 - Lookup: Stores indexed lookup list for vendors and clients.

Figure 1-80. Components – Datacap Database Server

Database Server

- One or more database servers house the Datacap databases.

Supported relational databases

The following relational databases are supported and can be used with Datacap.

- Oracle Versions 11j
- Microsoft SQL Server 2008 R2 or 2012
- DB2 10.5
- Microsoft Access
 - Is the default database that is used when a new application is created.
 - It should be used in test environments.
 - You can switch to using one of the other supported databases at any time.
 - You must switch to one of the more robust relational databases before going into production.

Four primary Datacap application databases are:

- Admin: Stores information about workflows, users, and groups.
- Engine: Stores batch information and statistics.
- Fingerprint: Stores page template data.
- Lookup: Stores indexed lookup list for vendors, clients, and products.

Components – Maintenance Manager Server

- Is a Datacap application.
- It automates administrative functions.
- It writes status messages to:
 - Internal log files.
 - Rulerunner log files.
- Uses logging actions to write information to the Maintenance Manager and windows log files.
- Sends emails that contain the internal log file.

Figure 1-81. Components – Maintenance Manager Server

Maintenance Manager

- Provides application monitoring and notification capabilities that can automate administrative functions such as resetting batches or archiving old batches.

Logging actions

Use the Logging actions to write information to the Maintenance Manager and Windows log files and to send emails that contain the internal log file.

During rule execution, Maintenance Manager writes status messages to an internal log file and the Rulerunner log file.

- The internal log file is maintained in memory and is used by theSendEmail action.
- The Rulerunner log file is stored in the application_name > batches > Maintenance Manager folder.
 - LogClear - Clears current Log.
 - LogConfigure - Configures features of aTM logging.
 - LogSendEmail - Sends email with log to comma-separated list of recipients.
 - LogWriteEventLog - Writes a message to the Event Log.

- LogWriteRecordSet - Outputs the results of ProcessRunSqlQuery to the error log.
- LogWriteSQLQuery - Outputs the constructed SQL query to the error log.

Components – Rulerunner Server

- Is a unified Windows service that runs background tasks that do not require operator interaction.
- Rulerunner tasks are configured in:
 - Application Manager, Rulerunner tab.
 - Rulerunner Manager, Workflow Job: Task tab.
- Scales vertically with multi-threading.
 - With Rulerunner Enterprise license, run as many threads as processors. Example: Quad processor four threads.
- Scales horizontally and does load-balance.
 - Add more servers
- Fingerprint Service.
 - Pattern matching service

Figure 1-82. Components – Rulerunner Server

Rulerunner task examples

The following tasks do not require operator intervention and therefore, can be run as background tasks under in the Rulerunner Server:

- Import – files, email*, fax*
- File format conversions
- Image enhancement and manipulation
- Classification
- Document assembly
- Recognition
- Export

Scaling to compensate for load

The background tasks that run in Rulerunner are where most of the heavy processing is done. There are two ways in which the Rulerunner load is distributed so that it does not become a bottleneck.

- Vertical scaling

- Is accomplished by configuring multiple Rulerunner threads in the Rulerunner Server.
 - The server must be a multi-processor system to effectively handle multiple threads.
 - You can achieve best performance by configuring as many threads as processors.
 - Example: four processor 4 threads
 - Maximum suggested processor to thread configuration is 150%.
 - Example: four processors 6 threads
 - Multi-threading does not require the purchase of more hardware.
 - Multi-threading does require an Enterprise Datacap license.
- Horizontal scaling
 - Is accomplished by configuring multiple identically configured servers in parallel in server farm.
 - The active – active server farm inherently distributed the load evenly across the available servers.

Fingerprint Service

In the component distribution sample that is used in this lesson, the fingerprint services are cohosted on the Rulerunner server. In other scenarios, the system load might require that the Fingerprint service is hosted on a dedicated fingerprint server.

- Pattern matching service – typically used when the number fingerprints is greater than 1000 fingerprints.

Components – File Server

- Stores work-in-progress batches
 - Images
 - Extracted data
 - Control files
- Stores configuration files that are used by all of the services
 - Application configuration files
 - Fingerprint patterns

Figure 1-83. Components – File Server

File server

- Applications are installed on the file server.
- Batches folder stores files that are generated during batch processing.

Components – External Systems

- External Databases.
 - Lookup Databases.
 - Export Databases.
- WebSphere for IBM services.
 - IBM Repository support.
 - IBM FileNet Content Manager.
 - IBM Content Manager.
 - IBM Content Navigator.
- CMIS compatible repositories.
(Content Management Interoperability Services).
- Other third party applications or services.

Figure 1-84. Components – External Systems

External Systems Server

There might be more than one external system server in the Datacap system.

External Databases

- Lookup Databases

There are a myriad of public lookup databases that can be accessed for example for company addresses, postal codes, telephone dial codes, services by company and more.

- Export Databases

It is possible that Datacap export documents might not go to a Content Management system; but instead the export documents are written to an external database.

WebSphere for IBM or third-party services

Is IBM's Web services server.

- IBM Repository support
 - IBM FileNet Content Manager & IBM Content Manager

The IBM available Content Manager systems

- IBM CMIS (Content Management Interoperability Services)
 - Is an open standard that CMIS Client uses to enable communication between Datacap applications and content management systems over the internet.
- IBM Content Navigator
 - IBM's web-based desktop that provides visibility into in the IBM Content Management systems, Case Manager system for automating business processes, IBM system configuration interfaces.
- Datacap Navigator
 - Is an IBM Content Navigator plug-in that provides web-based access and visibility into Datacap configuration, maintenance, and process interfaces.

Other third party applications or services might be:

- Hosted as applications installed directly on the External server or
- As web services that are implemented as WebSphere applications that provide connectivity to external system.

Datacap Folders for Services and Client Code

- Services folders
 - Taskmaster – Datacap Server Service
 - RRS – Repository for action libraries and global rulesets.
- Client Folders
 - DcDesktop - Datacap Desktop
 - DStudio – Datacap Studio Development tool for configuring and testing applications.
 - FastDoc – Rapid Application Development tool (RAD) and streamlined client.
 - tmweb.net – Datacap Web server component.
 - Tmweb.java – Datacap Navigator plug-in JAR file.
 - RV2 –Web application for report viewer.
 - wTM – Datacap Web Services REST API.

Figure 1-85. Datacap Folders for Services and Client Code

The default location for Datacap Capture installation on a Microsoft Windows system is in a folder that is named Datacap in the root of the C drive.

Datacap folder – Executive service.

- The subfolders of the Datacap folder are for:
 - Server Folders
 - Taskmaster folder
 - Taskmaster Server
 - Web Services when hosted as a windows service
 - Datacap Maintenance Manager (NENU)
 - Client Folders
 - Taskmaster folder
 - Web Services when hosting as a windows service
 - The DCDesktop folder
 - The FastDoc folder

- The tmweb.net and tmweb.java folders.
- The wTM folder is the web service for IIS hosting
- Utilities Folders
 - Taskmaster Folder
 - Datacap Maintenance Manager (NENU)
 - The support folder
 - Database copy utility
 - Database scripts, license Config
 - Web Server and Client setup utilities
- Data structure folders
- Applications
 - Each application has its own folder under the Datacap

Examples:

Built-in applications: APT, Flex, Medical Claims, TravelDocs

Class applications: Expense8, ExpenseDemo, FastStart, FastForm

Other Datacap Folders

- Utility folders
 - Support – Utilities and scripts for database creation, migration, and web server and client configuration.
- Miscellaneous support folders
 - dcshared – Datacap shared .dll, class, and other data and configuration files.
 - FingerprintService – The fingerprint web service application files.
 - license – Program license agreements.

Figure 1-86. Other Datacap Folders

Datacap Folders for Various Code and Data

These folders are the repositories for the code and data for Datacap support utilities.

The default location for Datacap Capture installation on a Microsoft Windows system is in a folder that is named Datacap in the root of the C drive.

- Datacap folder – Executive service
- The subfolders of the Datacap folder are for:
 - Utilities process folders.
 - Data structure folders

The Datacap.xml file

- This file is the directory of all the applications in a Datacap environment.
- It contains:
 - The Datacap version number
 - The name and folder location of each application in the environment.

```

1  <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2  <datacap ver="9.0">
3      <app name="FieldTest" ref="FieldTest"></app>
4      <app name="MyFormApp" ref="C:\Datacap\MyFormApp"></app>
5      <app name="MyCMISApp" ref="C:\Datacap\MyCMISApp"></app>
6      <app name="Flex" ref="Flex"></app>
7      <app name="Forms" ref="Forms"></app>
8      <app name="Learning" ref="Learning"></app>
9      <app name="TravelDocs" ref="TravelDocs"></app>
10     <app name="APT" ref="APT"></app>
11     <app name="Medical Claims" ref="Medical Claims"></app>
12     <app name="FormTemplate" ref="Templates\FormTemplate"></app>
13     <app name="LearningTemplate" ref="Templates\LearningTemplate"></app>
14 </datacap>

```

Introduction to Datacap © Copyright IBM Corporation 2016

Figure 1-87. The Datacap.xml file

Help path

- Datacap> Datacap 9.0.1 > Datacap application development > Creating a Datacap Maintenance Manager application > Updating the datacap.xml file on the Datacap server
- Datacap > Datacap 9.0.1 > Installing > Installing and configuring in a client/server environment > Datacap installation and configuration in a client/server environment > Complete the Datacap Web Client server configuration > Setting the location of the datacap.xml file

The screen capture shows a Sample datacap.xml file.

The datacap.xml file

- The datacap.xml file is in the Datacap root folder. It contains the Datacap version number and a tag that defines the name for each of the defined Datacap Capture applications.
- The version and exact build number are also shown in the C:\Datacap\version.txt file.

Important: The Datacap Application Service is case-sensitive. When you add or change entries in the datacap.xml file, make sure that the case matches the case of the UNC paths, folders, and file names.

- The ref attribute of the app node in datacap.xml uses a relative path if no folder is specified.

Review question (1)

In the table in the notes area, enter the number that corresponds to the Datacap component name from the following Datacap system architecture diagram.

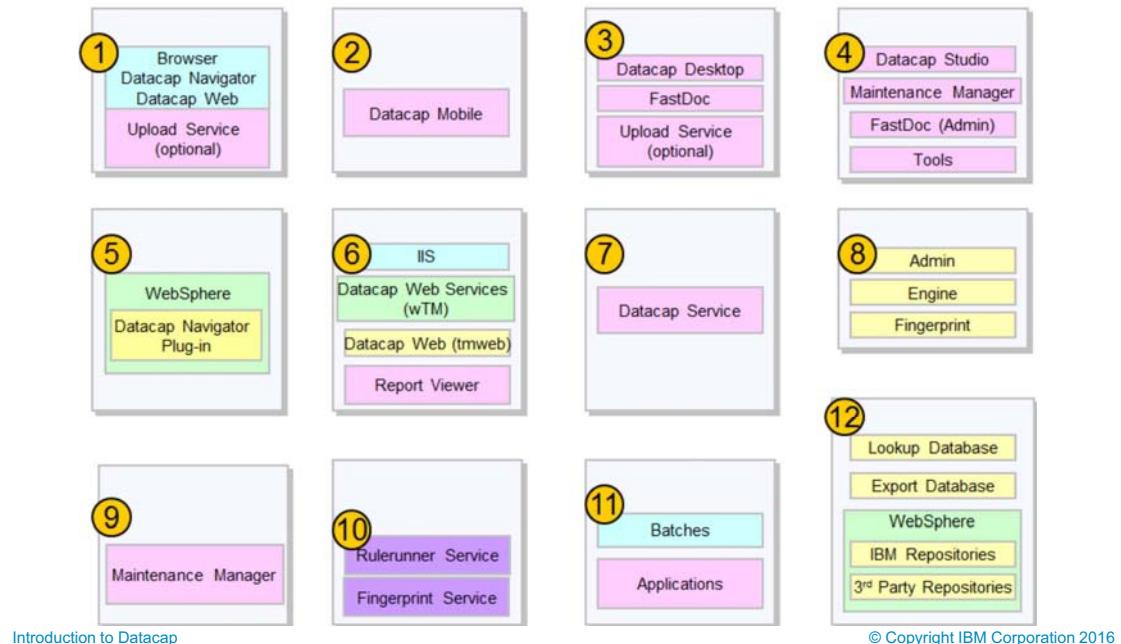


Figure 1-88. Review question (1)

Write your answers here:

Datacap component name Component Number

Developer Workstation

Web Client Workstation

Datacap Navigator Server

Datacap Web Server

Database Server

Maintenance Manager Server

Datacap Client Workstation

Mobile Device

File Server

Datacap Server

Rulerunner Server and Fingerprint Services Server

External Systems



Review answers (1)

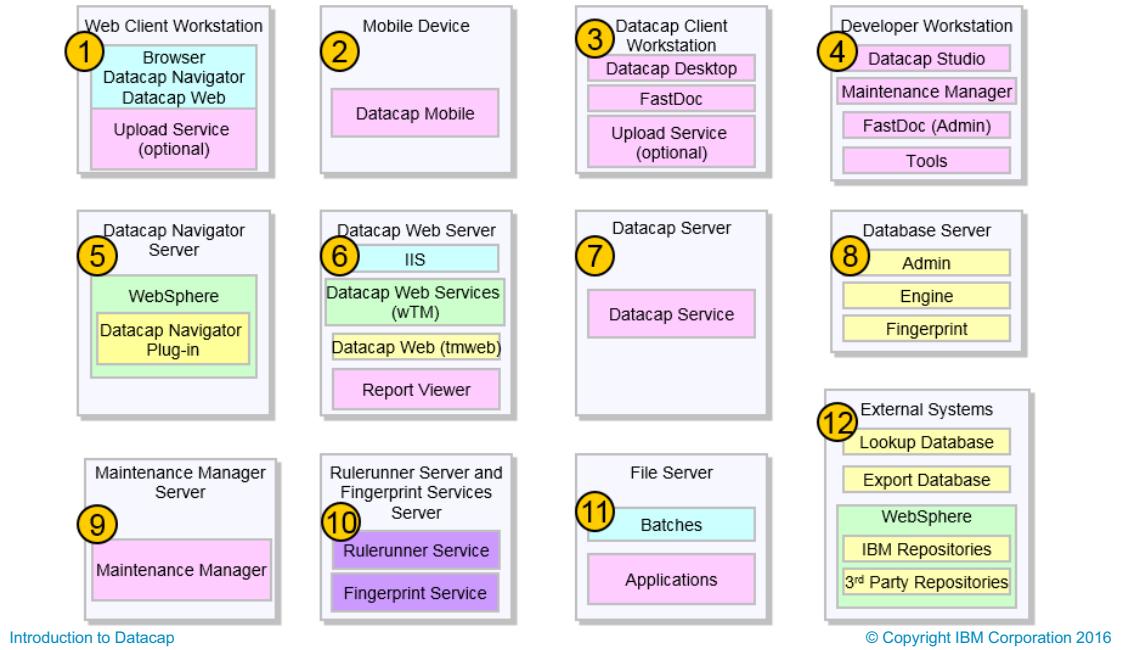
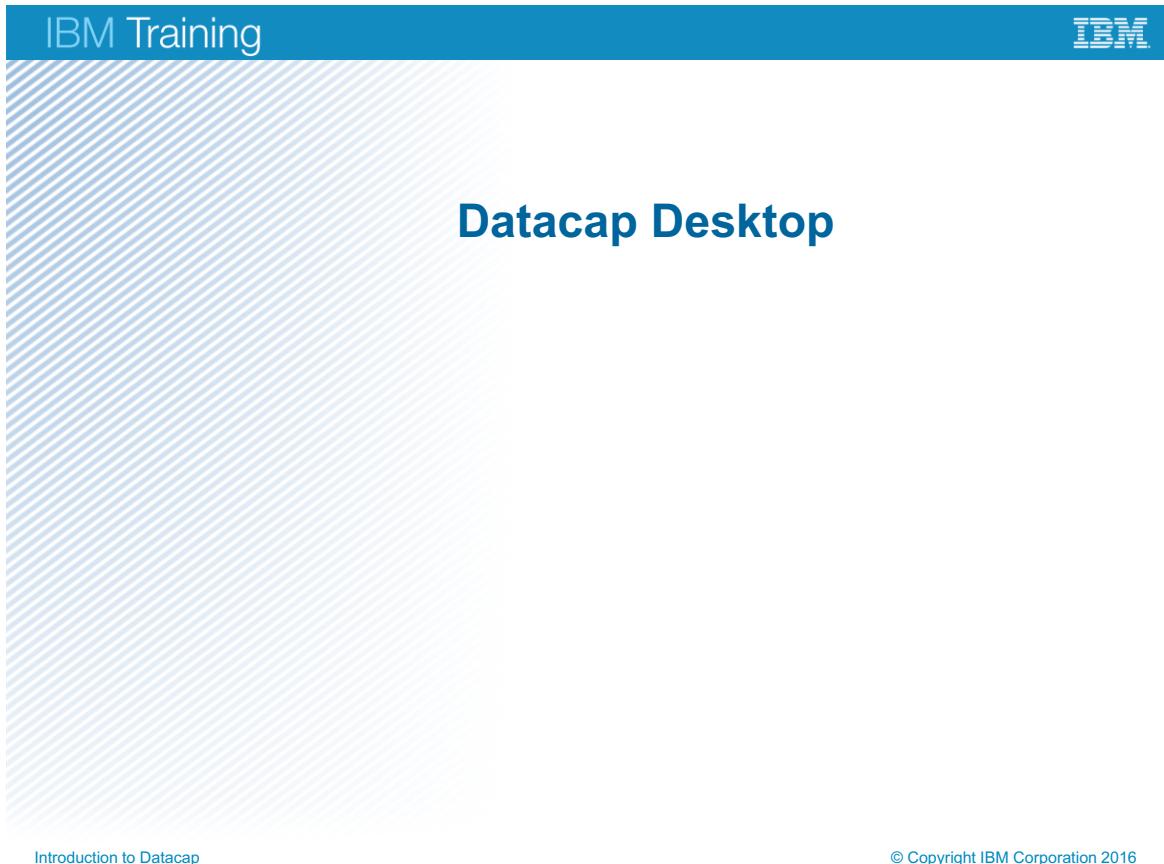


Figure 1-89. Review answers (1)

The answers are:

<u>Datacap component name</u>	<u>Component Number</u>
<u>Developer Workstation</u>	(4).
<u>Web Client Workstation</u>	(1).
<u>Datacap Navigator Server</u>	(5).
<u>Datacap Web Server</u>	(6).
<u>Database Server</u>	(8).
<u>Maintenance Manager Server</u>	(9).
<u>Datacap Client Workstation</u>	(3).
<u>Mobile Device</u>	(2).
<u>File Server</u>	(11).
<u>Datacap Server</u>	(7).
<u>Rulerunner Server Fingerprint Services Server</u>	(10).
<u>External Systems</u>	(12).

Lesson 1.6. Datacap Desktop



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-90. Datacap Desktop

Lessons

- Datacap overview
- Datacap process
- Role-based Datacap clients
- Architecture configurations
- Architecture components
-  Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-91. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you create applications that are processed with the Datacap Desktop Client.
- To do these tasks effectively, you must be familiar with the Datacap Desktop client.

Figure 1-92. Why is this lesson important to you?

Capabilities Overview

- Use Datacap Desktop client to do the following tasks:
 - Create batches, identify pages, create documents, and export batches.
- To use Datacap Desktop to complete a task:
 - Must specify Desktop as the program to use in setup page for the task.
- The Datacap Desktop:
 - Has a single login point for completing tasks.
 - Can select any application that is configured in the Datacap system for processing tasks.
 - Includes the Queue Monitor in which users with appropriate privileges can view or start pending tasks for the selected application.

Figure 1-93. Capabilities Overview

Create batches:

- Is the process of scanning batches of paper documents or accessing a folder of electronic documents and queuing them for further processing.

Identify pages:

- Identify a page by the layout or identifiable fields so that further processing can be done to extract and verify the data on the page.

Create Documents:

- Group pages into documents and create a document that represents the grouped pages.

Export batches

- Batches of processed documents are sent on the next step in the business process. The next step might be to export the documents to a repository where they are stored and can be searched, sorted, and categorized for ease of access. The next step might also be direct them to subsequent application for further processing.

Use Datacap Desktop with your Datacap applications

- Use Datacap Desktop to complete the following tasks:
- VScan
 - Imports files from a file system.
- Scan
 - Imports physical paper documents by scanning.
- PageID
 - Identifies the page type of a scanned image.
- Profiler
 - Creates documents, extracts and validates data and routes documents to the next step.
- Verify/Fixup
 - A user can verify, and correct document and extracted data values.
- Export
 - Exports batches to a specified location.

Figure 1-94. Use Datacap Desktop with your Datacap applications

VScan

- This task is for a virtual scan that imports files from a specified location, and is used mostly for demonstration purposes.

Scan

- In production environments, a scan task is configured to scan paper documents that use either TWAIN or ISIS scanners.

Profiler

- Locates and extracts data values from pages.
- Arranges identified pages into documents.
- Does confidence testing on all characters read from the page.
- Does validation on extracted field values.
- Optionally routes documents to the next task based on validation results.

The documents that are error-free can be routed directly to the Export task. The documents that have errors are routed to an operator to verify and correct potential error.

Verify/Fixup

- This task requires user input to correct any errors or integrity issues that a preceding task encounters.
- The batch does not continue to the next task until the Fixup task is completed.

Background

- This task completes all tasks that do not require user intervention, including PageID, Profiler, and Export, and automates the completion of pending batches.

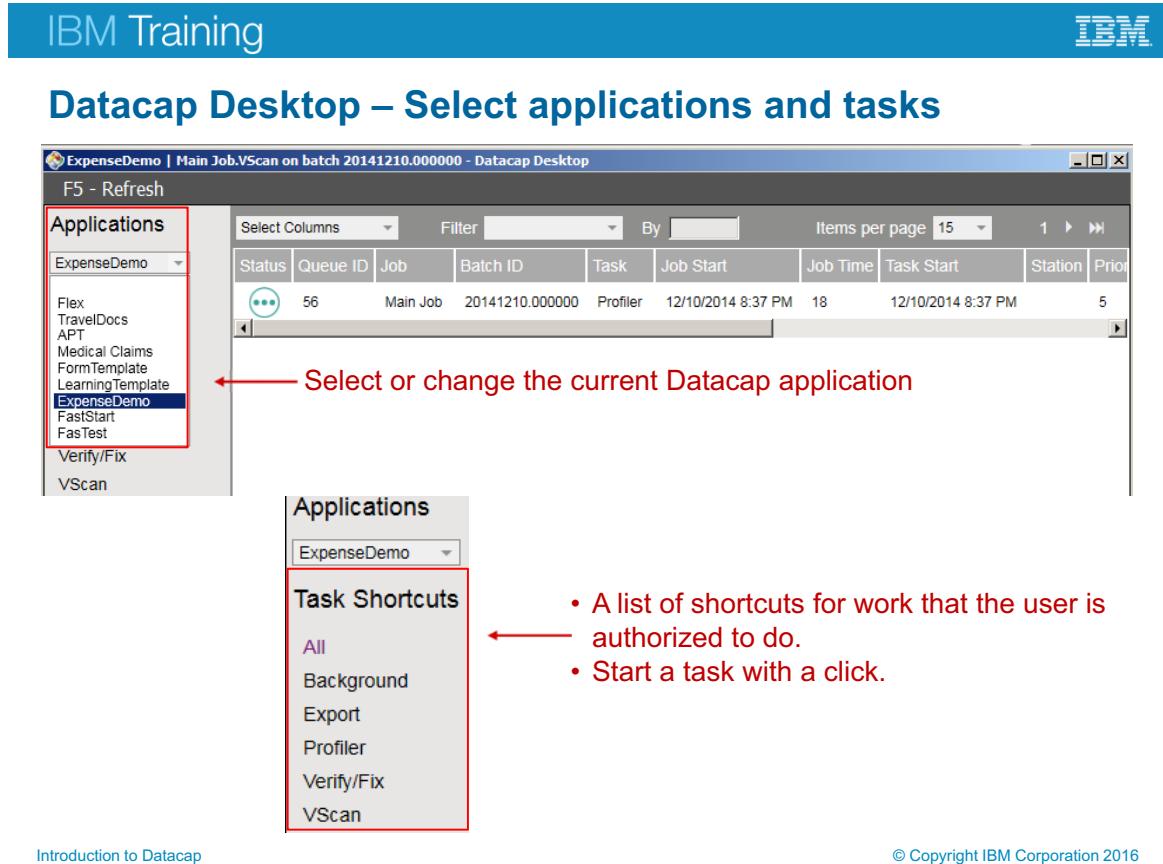


Figure 1-95. Datacap Desktop – Select applications and tasks

Desktop allows the users to switch between applications, and tasks with ease.

- Users can select an application or change to another application from the list.
- After selecting a particular application, a list of allowed shortcuts for work that the user is authorized to do is displayed below the application.
- When the user clicks a shortcut to start, what gets displayed in the view on the right pane depends on:
 - The user permissions
 - Application and Task Shortcut selections
 - The mode (manual, automatic, or manual for hold) of the task.
- Desktop tracks the last application and that shortcut that the user was working on, and activates them on the next launch.

Job Monitor view

- If user has Job Monitor view privilege, the "All" shortcut is listed.
- When the user selects "All", the Job Monitor view is shown in the right pane.
- The view lists the batches for all the allowed job-task combinations for the user.

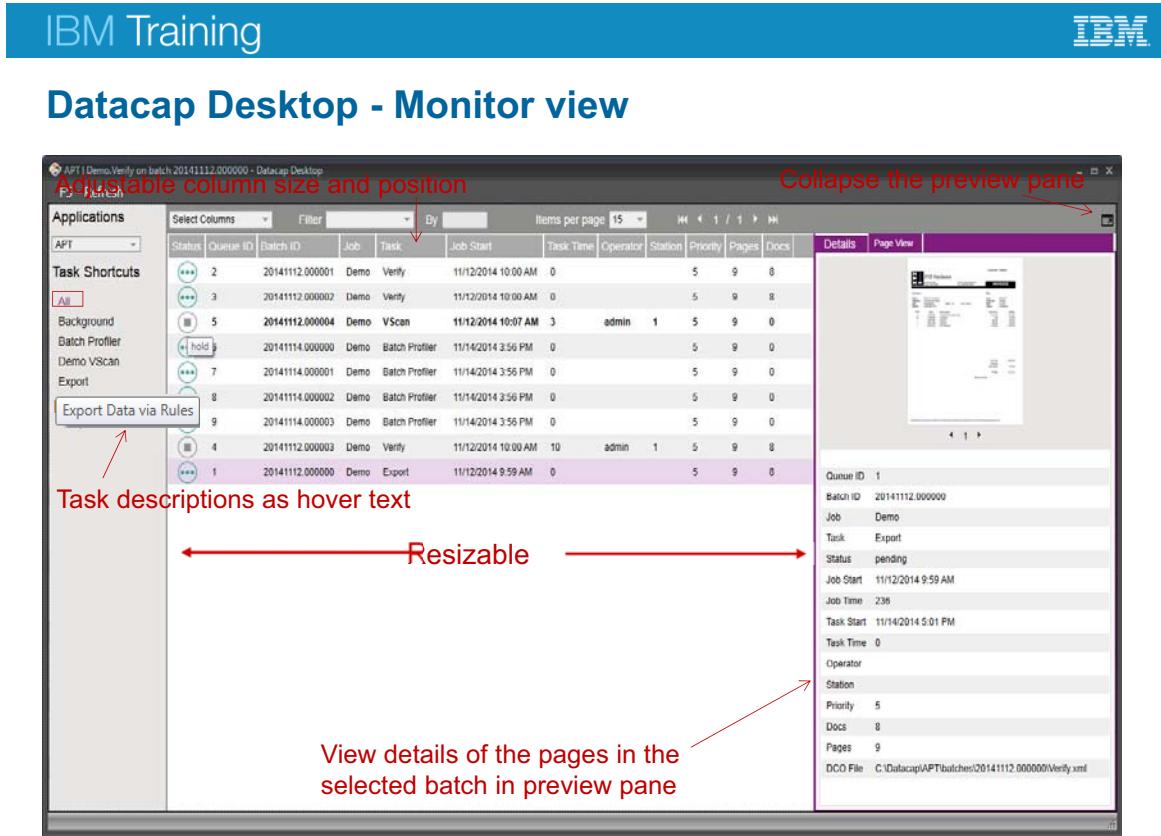


Figure 1-96. Datacap Desktop - Monitor view

Batches in Monitor view

- The status of a batch is shown as an icon in the first column. The active task has a different color that is compared to the inactive tasks in the list.
- When you hover over a task, description for that task is shown.
- Double-click of a batch starts the associated task.

Task Preview

- Single click of a batch shows its details in the preview pane on the right.
- The preview has two tabs: Details and Page View.
 - Details View shows the batch details such as the number of pages and documents.
 - On the Page View, you can look at the pages in a batch without having to start a task.
- Preview pane is not shown when no batch is selected. You can resize and collapse the preview pane.

Column size and position

- Drag the individual column by header to change its relative position in the table. Click a column header to sort by that column.



Select columns and apply filter in the Monitor view

F5 - Refresh

Applications		Select Columns														
ExpenseDemo		<input checked="" type="checkbox"/> Queue ID	<input checked="" type="checkbox"/> Batch ID	<input checked="" type="checkbox"/> Job	<input checked="" type="checkbox"/> Task	<input checked="" type="checkbox"/> Status	<input checked="" type="checkbox"/> Job Start	<input checked="" type="checkbox"/> Job Time	<input checked="" type="checkbox"/> Task Start	<input type="checkbox"/> Task Time	<input type="checkbox"/> Operator	<input checked="" type="checkbox"/> Station	<input checked="" type="checkbox"/> Priority	<input checked="" type="checkbox"/> Docs	<input checked="" type="checkbox"/> Pages	<input checked="" type="checkbox"/> DCO File
Task Shortcuts																
All																
Background																
Export																
Profiler																
Verify/Fix																
VScan																

← You can select or clear the columns to display

Apply filter Specify # Page Navigator

Select Columns		Filter		Task		By		Verify		Items per page		15		Page Navigator	
Status	Queue ID	Batch ID	Job	Task	Job Start	Task Time	Operator	Station	Priority	Pages	Docs				
(•)	2	20141112.000001	Demo	Verify	11/12/2014 10:00 AM	0			5	9	8				
(■)	4	20141112.000003	Demo	Verify	11/12/2014 10:00 AM	10	admin	1	5	9	8				
(■)	3	20141112.000002	Demo	Verify	11/12/2014 10:00 AM	88	admin	1	5	9	8				

Introduction to Datacap © Copyright IBM Corporation 2016

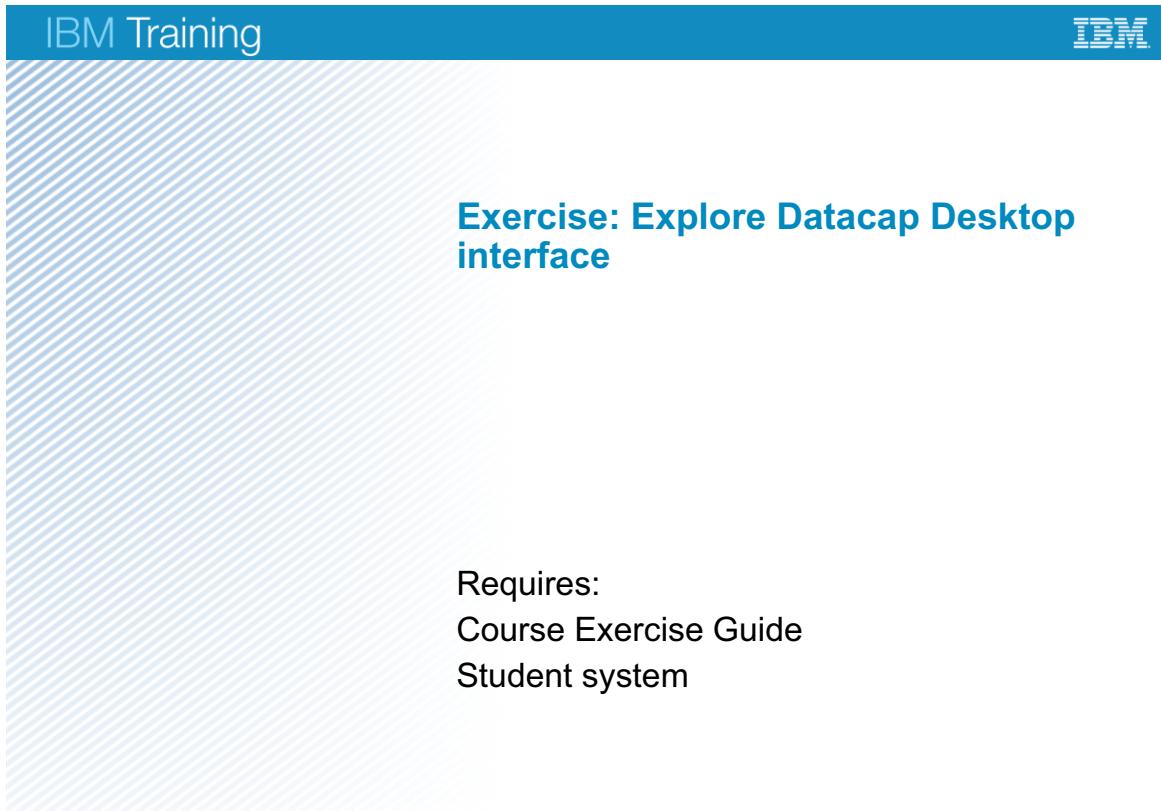
Figure 1-97. Select columns and apply filter in the Monitor view

- You can select what columns to show or clear a column to hide.

- Apply Filter

Use “Filter-By” to view batches based on a criterion. For example, Task = Verify as shown in the screen capture to list only the Verify tasks.

- You can specify the number of items (rows) per page that you want to show.
- Use the page navigator to move through pages.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-98. *Exercise: Explore Datacap Desktop interface*

Exercise objectives

- Explore Datacap Desktop interface.



Figure 1-99. Exercise objectives

Lesson 1.7. Application design

Application design

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-100. Application design

Lessons

- Datacap overview
- Datacap process
- Role-based Datacap clients
- Architecture configurations
- Architecture components
- Datacap Desktop
-  Application design
 - Introduction to Datacap Navigator
 - Datacap web client (tmweb)

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-101. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users.
- To do these tasks effectively, you must be familiar with the design principles for the Datacap Application.

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-102. Why is this lesson important to you?

Scenario for the complete application

- Your company needs to enter data from expense documents and you want to automate the process with Datacap.
- The application processes four image types and assembles them into three document types.
 1. Car (Document type)
 - Rental_Agreement (Page type)
 2. Air (Document type)
 - Air_Receipt (Page type)
 3. HR (Document type)
 - HR_Page (Page type)
 - Donation_Receipt (Page type)

Figure 1-103. Scenario for the complete application

Scenario for the complete application

- Your company has numerous external sales, and technical support personnel who are continually traveling between customer sites. So, they have many car rental documents and airline ticket receipts. In addition, the project processes humanitarian relief fund expenses that are not related to travel with their expense reports.
- Your company authorizes three car rental companies, and three airlines for company travel expenses. In some locations, employees are allowed to use other car rental and airline companies. You are going to create a Datacap application that recognizes car rental agreements and air ticket receipts for the preferred vendors.
- Your application also can process unfamiliar documents when they are encountered. The application must be able to recognize name, date of service, and total cost of service charged.
- The final export data is stored as documents in the IBM FileNet Content Manager repository.
- The completed project (“ExpenseDemo”) that is based on this scenario is installed on your class image.
- You are going to scan and process a batch with this application in this lesson.

Steps to design and configure an application

- Do a complete end-to-end evaluation of the process to be automated.
- Analyze and define the requirements for the batch process.
- Analyze the data source.
- Implement a sandbox system.
- Configure the Datacap application.
- Test and adjust the application.

Figure 1-104. Steps to design and configure an application

Do a complete end-to-end evaluation

- Analyze the end-to-end business process that is associated with your corporate documents that need to be captured.
- Determine of what data is going to be captured so you know what to look for in the documents.
 - How are the images to be captured including resolution and color properties of the images.
 - How to locate and extract the data?
 - How to validate the data?
 - Where is the output to be sent?
 - Does it need any special format?
- Determine the steps people must do.
- At what point can the documents be removed from the system or archived?

Figure 1-105. Do a complete end-to-end evaluation

Analyze and define the requirements

- For your Expense Report claims, extract pertinent fields data and validate:
 - Employee names or numbers.
 - Service vendor names or codes.
 - Amount of the expense.
- Is there a need for extra documents and pieces of information later like:
 - Accident reports.
 - Damage quotations.
 - Baggage loss claims.

Figure 1-106. Analyze and define the requirements

Analyze the input documents

- What channels of input are supplying the images or other files?
- What resolution and color properties will the images have?
 - 200 DPI black and white unless otherwise specified.
- How can the images be classified (assign page types)
 - Fingerprints, regular expressions, graphical patterns (logos), Content Classification, and so on.
- How can the data elements be on each document.
 - Zone coordinates, regular expressions, locate actions.
 - Are all the required data elements on the documents?
- How can documents be separated in the stream of input images?
 - Separator sheets (barcodes), page types, manually.

Figure 1-107. Analyze the input documents

Implement a sandbox system

- Install full Datacap installation on an all-in-one system for the development sandbox.
- Configure the Datacap application – database, workflow, document hierarchy, rules.
- Configure the content repository or target system, and lookup databases if any, if not already done.
- Test the system end to end with one document type.
- Test the remaining document types and all functions.

Figure 1-108. Implement a sandbox system

Configure Datacap Application

- Configure Datacap to capture the documents and metadata that you expect. These tasks include:
 - Create a Datacap application.
 - Create the document hierarchy, the data fields, and zones.
 - Configure tasks, Rulesets, the lookups, database feeds.
 - Design screens and dialogs.
 - Configure functional security.
 - Configure export of the documents and the metadata.
 - Configure the reporting, activity monitoring, and notification.

Figure 1-109. Configure Datacap Application

Application development strategy

- Create new Application with Application Wizard
 - Application Wizard can be called from Datacap Studio and FastDoc.
 - Use the Application Wizard icon on the toolbar to create an application with one of the following methods:
 - Use Form or Learning template.
 - Copy an existing application.
- Do initial development in FastDoc
 - Configure Batch Structure
 - Configure Rulesets
 - Define fingerprints
 - Test profiles and rules
- Do advanced development in Datacap Studio
 - Configure conventional rulesets
 - Develop advanced custom rulesets



Figure 1-110. Application development strategy

Create new Application with Application Wizard

The Application Wizard Gives you a head start on application development by generating a basic application framework.

- Copy an application – What do you get with this option?
 - A copy of a completely functioning application as a starting point.
 - You can modify and customize it to meet the functional requirements for your new project.
- Create an RRS application – What do you get with this option?
 - An option of selecting one of two built-in application templates as a starting point for your new application.
 - A complete application folder structure.
 - A skeleton Document hierarchy (DCO), which you can expand to provide the document structure for your planned document batches.
 - A functional set of built-in compiled rulesets.
 - A set of workflow task profiles, preconfigured to run data capture tasks.
 - A custom mapping configuration to map rules from the rulesets to the DCO objects.

Note:

- a. You can create and save your own custom templates to use for future projects.
- b. When creating new DCO objects, you can configure them to inherit rulesets from existing objects of the same type.

Do initial development in FastDoc

FastDoc provides a quick and intuitive way to do the following basic application development steps.

- Add document hierarchy components.
 - Add documents
 - Add pages to a document
 - Add fields to pages
- Define fingerprints
 - Add an image that represents the page fingerprint.
 - Define fingerprint classes.
 - Define fingerprint types.
 - Define field zones for each field you want to extract from the page with zonal recognition.
- Configure Ruleset settings for UI compiled rulesets
 - Batch – Import files location and file type
 - Batch – Identify Pages
 - Batch – Create Documents
 - Batch – Convert Files To Images
 - Page – Image Enhancement
 - Page – Recognize Pages and Fields
 - Field – Recognize Pages and Fields
 - Field – Validate fields
- Test the profile and ruleset configuration.
 - Test that documents and fields are properly identified.
 - Test that field values are extracted correctly.

Do advanced development in Datacap Studio

- Configure conventional rulesets that do not have a UI, which enables them to be configured in FastDoc.
- Develop advanced custom rulesets for doing things that are not possible with the built-in compiled rulesets.

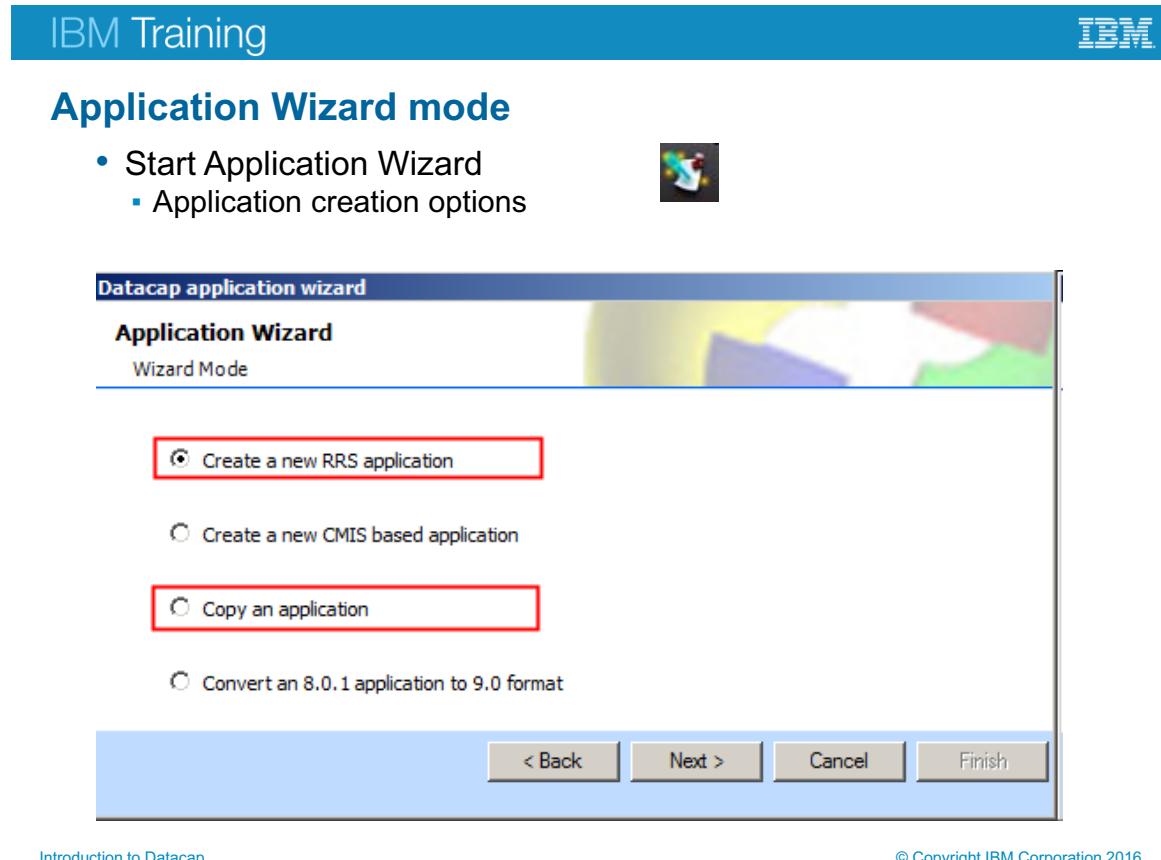


Figure 1-111. Application Wizard mode

Start Application Wizard

Start either Datacap Studio or FastDoc (Admin)

1. Click Start > All Programs > IBM Datacap Developer Tools > Datacap Studio and click close.
Click Start > All Programs > IBM Datacap Developer Tools > FastDoc (Admin) Click Local and click Login.
2. Click the Application Wizard icon on the toolbar of either interface and click Next.

Application creation options.

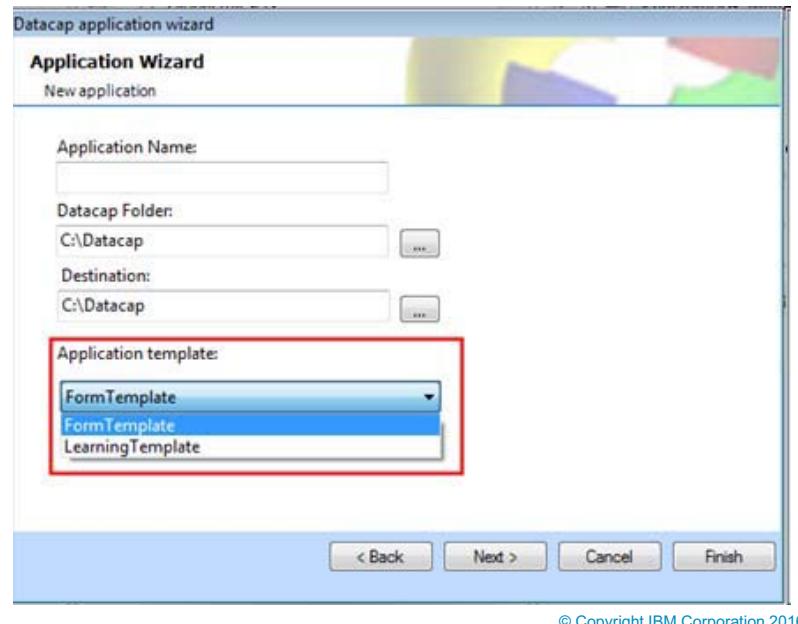
- Create an RRS application
 - Use this option to create a new allocation skeleton from one of the built-in application templates.
- Create a CMIS-based application
 - Use this option to create a new allocation skeleton from one of the built-in application templates.
 - Configured to use CMIS (Common Management information Service) connectors for backend services.

- Copy an application
 - Use this option if you already have an application that is very close to the configuration that you need for your new project.
- Convert an 8.0.1 application to 9.0 format
 - Use the option if you have an old application that was developed in Datacap 8.0.1 that needs to be brought up to Datacap 9.0 format.



Application Wizard New Application window

- Application Templates
 - Datacap currently released with two built-in templates
 - Form
 - Learning



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-112. Application Wizard New Application window

Application Template options.

Forms Template

- Click Forms to use the application for structured images.
- When you know the types of data that you want to capture and where that data is on each image, select the Forms application.
- For example, a 1040EZ tax form and the types of data on the form, such as name and address, are in the same location on every 1040EZ form.
- The Forms application sets up a workflow that you can match against your fingerprints.

Learning Template

- Click Learning to use the application for unstructured images.
- Select the Learning application when you:
 - Know the types of data that you want to capture but
 - Do not know where that data is contained in the image (because the location of the data is different on each image).

- For example, if you want to capture the date, amount, and tax for expenses from different hotels, the receipt images from each hotel are unique.
- The location of the data that you want to capture differs for each hotel receipt image so the data cannot be identified with Datacap fingerprints.
- The Learning application template sets up a workflow where you can add rules, such as Locate rules, for Datacap to learn the different hotel receipt formats as they are encountered.

Test and adjust

- Before the system can be put into production, the following tasks and issues must be addressed:
 - Testing and adjustments under various load conditions.
 - Bottlenecks that occur between the various system components.
 - Identify bottlenecks with reporting tools.
 - Example: Datacap Report Viewer

Figure 1-113. Test and adjust

Before the system can be put in production, the following tasks and issues must be addressed:

- Testing and adjustments under various load conditions are required to get to the expected results.
- Bottlenecks that require a remedy might occur between the various system components.
- Datacap has reporting tools to help identify bottlenecks.
 - Example: Datacap Report Viewer

Datacap Report Viewer

- Use the Reporting actions to write information to the report tables in the Engine database for use by Datacap Report Viewer.
- The Reporting actions can query the active users on an application and set the database tables that contain the reports for processed batches and users.

Best Practices

- Transfer relevant information into workflows.
- Make the post export workflow drive the business process.
- Normalize the data as early as possible in the process.
- Integrate with your business systems early in the process.
- Standardize types of documents and the forms.
- What is the fastest entry point in the system?
 - Is it at a local level?
 - Is it at the central level?
 - Who is best placed to supply the data needed?

Figure 1-114. Best Practices

IBM Advanced Document Imaging PIE

Best of breed capabilities in a single offering

New release of PIE - 5.2.1

Now includes:

- Rulerunner Enterprise
- Fingerprinting service



A single highly integrated platform!

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-115. IBM Advanced Document Imaging PIE

IBM Production Imaging Edition is a comprehensive solution for managing the entire Document imaging lifecycle. It:

- Helps organizations quickly and easily capture and extract key information from documents and convert it to images that are stored electronically
- Uses IBM FileNet Content Manager repository to improve manageability and flexibility
- Incorporates IBM Case Foundation components to help increase process performance and productivity
- Improves document usability with enterprise-wide document image viewing, annotation, and redaction

Product capabilities in Production Imaging Edition

- IBM Datacap for advanced Document capture
- IBM FileNet Content Manager for content management
- IBM Case Foundation components for business process management
- IBM Content Navigator and Daeja ViewONE Professional for image viewing, annotation, and redaction.

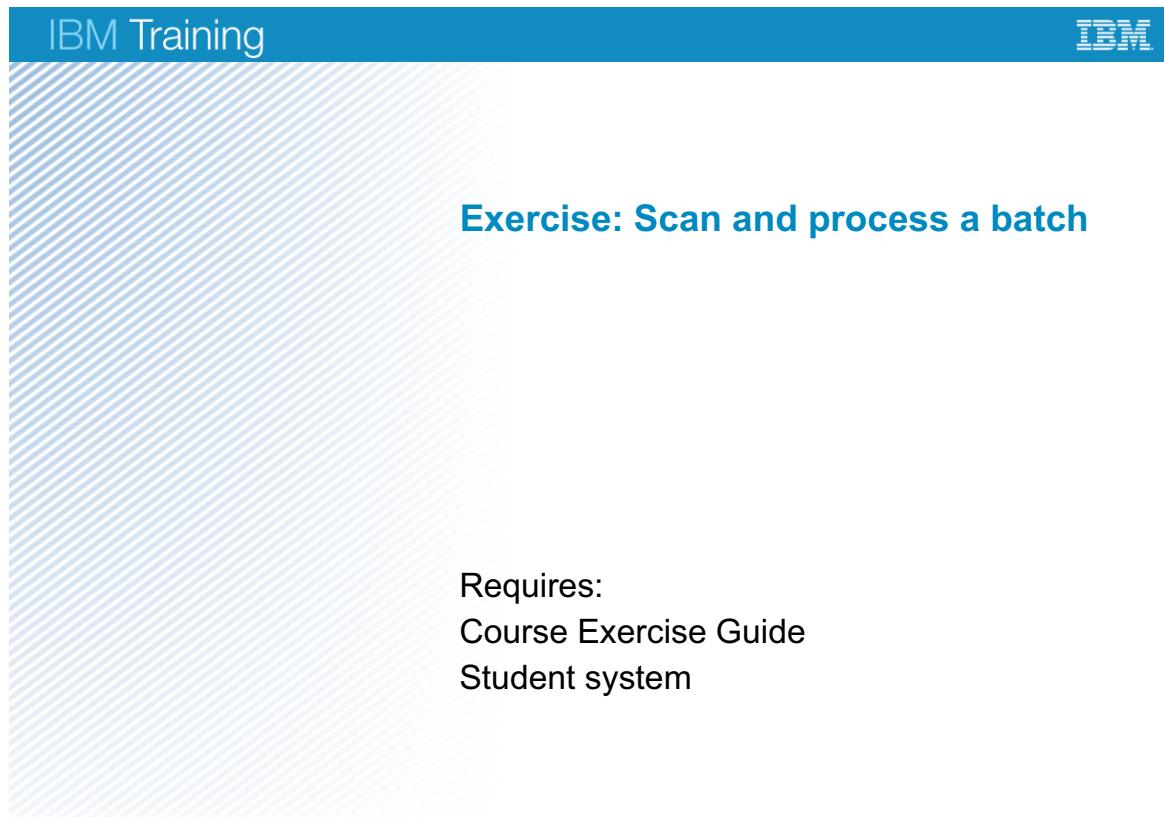


Figure 1-116. Exercise: Scan and process a batch

Exercise objectives

- Scan and process a batch.



Figure 1-117. Exercise objectives

Lesson 1.8. Introduction to Datacap Navigator

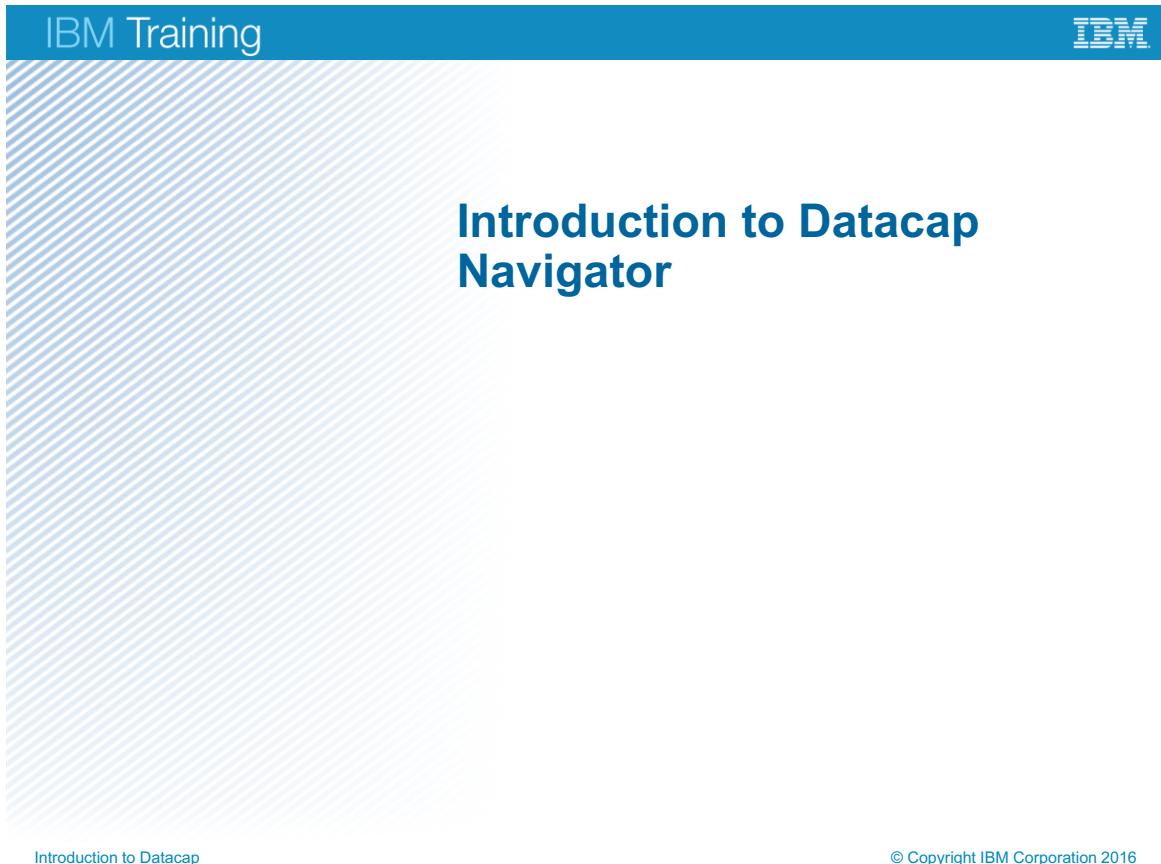


Figure 1-118. Introduction to Datacap Navigator

Lessons

- Datacap overview
 - Datacap process
 - Role-based Datacap clients
 - Architecture configurations
 - Architecture components
 - Datacap Desktop
 - Application design
-  [Introduction to Datacap Navigator](#)
- Datacap web client (tmweb)

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-119. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you test your application workflow tasks in Datacap Navigator and Datacap Mobile.
- As a business user, you scan and process your batches in Datacap Navigator and Datacap Mobile.
- To do these tasks effectively, you must be familiar with the Datacap Navigator interface and with processing batches.
- For Datacap Administrators you also use the Datacap Web Client (tmweb) to configure applications workflows, tasks, users, groups, and task shortcuts.

Figure 1-120. Why is this lesson important to you?

What is Datacap Navigator?

- IBM Content Navigator is a web-based client and framework.
 - To work with content from content management repositories.
- Datacap includes a plug-in, which operates within Content Navigator.
- Datacap Navigator provides a web user interface that:
 - Supports both user and administrative features for Datacap.
 - Provides access to the Datacap Job Monitor and to run tasks.
 - Enables you to add users and groups.
 - Provides access to configure stations and workflows.

Figure 1-121. What is Datacap Navigator?

- Datacap is combined with other Enterprise Content Management capabilities into a single user interface framework in Content Navigator.

Views in Datacap Navigator

- In Datacap Navigator, there are two views (Feature).
 - Datacap View for the business users.
 - Datacap Administration View for administrators.
- You can configure and customize Datacap Navigator for different roles.
 - Business Users (To scan, verify, and monitor jobs)
 - Administrators (To configure and administer)
- You can add more features such as Browse or Search.
 - To browse folders or search for documents that are exported.
 - To scan the documents directly from the IBM Content Navigator.

Figure 1-122. Views in Datacap Navigator

- Completed Datacap documents can be exported to a repository for storage. These repositories can be accessed in Content Navigator.
- The Content Management capabilities in Content Navigator open the Datacap documents in Search or Browse Views.
- The topic “Configure Datacap Navigator for different roles” is presented in another course.

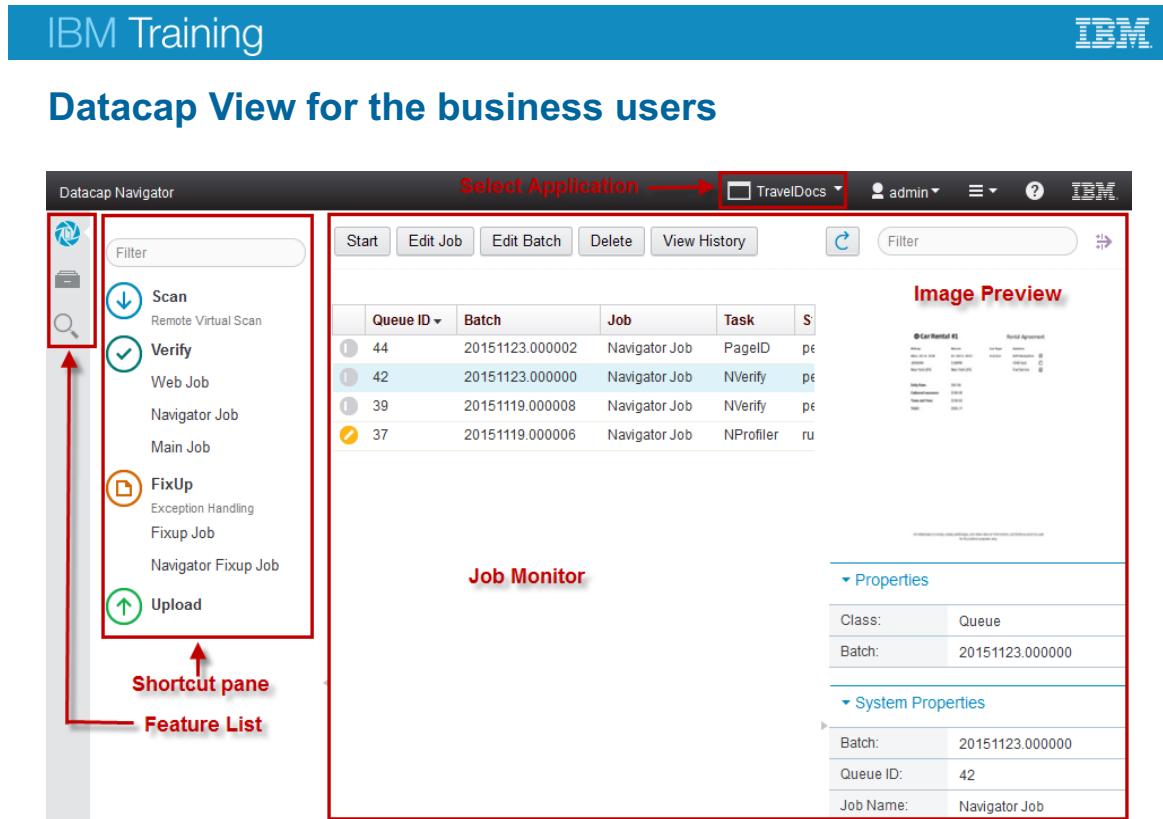


Figure 1-123. Datacap View for the business users

The Datacap view that is shown in the screen capture is for Datacap business users.

Feature List

- Datacap Navigator has two features by default (Datacap View and the Administrator View).
- You can optionally add Browse (folder cabinet icon) and Search (magnifying glass icon), which are part of IBM Content Navigator, as shown in the screen capture.
- Browse and Search views helps to view the documents that are exported from Datacap process to repositories.

Shortcut pane

- It contains a list of shortcuts for operations.
- The pane shows a list of all of the tasks that they are authorized to run.

Job Monitor

- The list of the batches that are in process and the completed ones.

Select Application

- The list of Datacap applications that are available are listed.

- Select an application and Job Monitor lists the batches for that application.
- You can monitor jobs for the batches that are processed in Datacap Desktop.

IBM Training

Job Monitor

Refresh admin ? IBM

Start Edit Job Edit Batch Delete View History ← Toolbar

Properties Filter

System Properties

Advanced Properties

Job List

Property panel

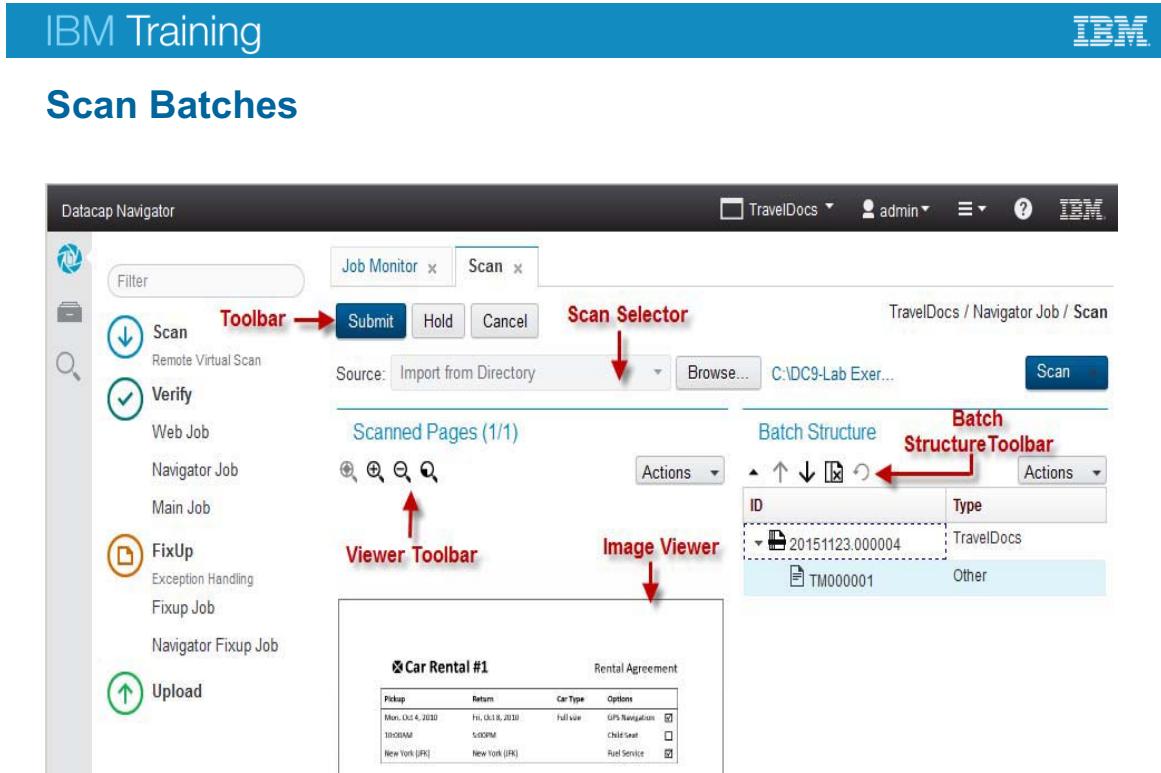
Queue ID	Batch	Job	Task	Status	Job Start
26	20141203.00001	Navigator Job	Scan	hold	12/3/2014, 3:23 P
25	20141203.00001	Navigator Job	Scan	running	12/3/2014, 3:21 P
24	20141203.00001	Navigator Job	Scan	hold	12/3/2014, 2:53 P
20	20141201.00001	Main Job	Verify	pending	12/1/2014, 5:50 P
12	20140904.00001	Navigator Job	NVerify	pending	9/4/2014, 11:53 A
11	20140904.00001	Navigator Job	NUpload	hold	9/4/2014, 8:45 AM
10	20140904.00001	Main Job	Verify	pending	9/4/2014, 8:35 AM
9	20140904.00001	Main Job	Verify	pending	9/4/2014, 8:34 AM
8	20140903.00001	Navigator Job	NVerify	hold	9/3/2014, 9:53 AM
7	20140901.00001	Navigator Job	Export	Job done	9/1/2014, 1:08 AM
6	20140901.00001	Navigator Job	NVerify	hold	9/1/2014, 1:00 AM
4	20140829.00001	Main Job	Verify	pending	8/29/2014, 3:35 P
3	20140829.00001	Main Job	Verify	pending	8/29/2014, 3:35 P
2	20140829.00001	Main Job	Verify	pending	8/29/2014, 1:44 P
1	20140829.00001	Main Job	Verify	hold	8/29/2014, 1:44 P

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-124. Job Monitor

- Job list
 - List all or filtered jobs
- Toolbar
 - Start a job, Edit Job, Edit Batch, Delete Batches, View History
- Refresh
 - Refresh job list
- Filter
 - Easily filter with quick search
 - Support both client and server end filter in advanced filter
- Job information pane
 - Show batch cover page thumbnail
 - Show job properties



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-125. Scan Batches

Help path

- Datacap > Datacap 9.0.1 > IBM Datacap V9.0.1 documentation > Accessibility features of Datacap Navigator
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/dcnav002.htm
- Toolbar
 - Submit Batch, Hold Batch, Cancel Batch
- Scanner selector
 - Lists all available scanners.
 - You can browse to local directory to get stored images.
- Viewer toolbar
 - Zoom In/Out/Quarter, Fit to Width/Height, Previous/Next Page, Show viewer in dual monitor
- Batch structure toolbar
 - Delete one or all pages, Move Up/Down to reorder the pages

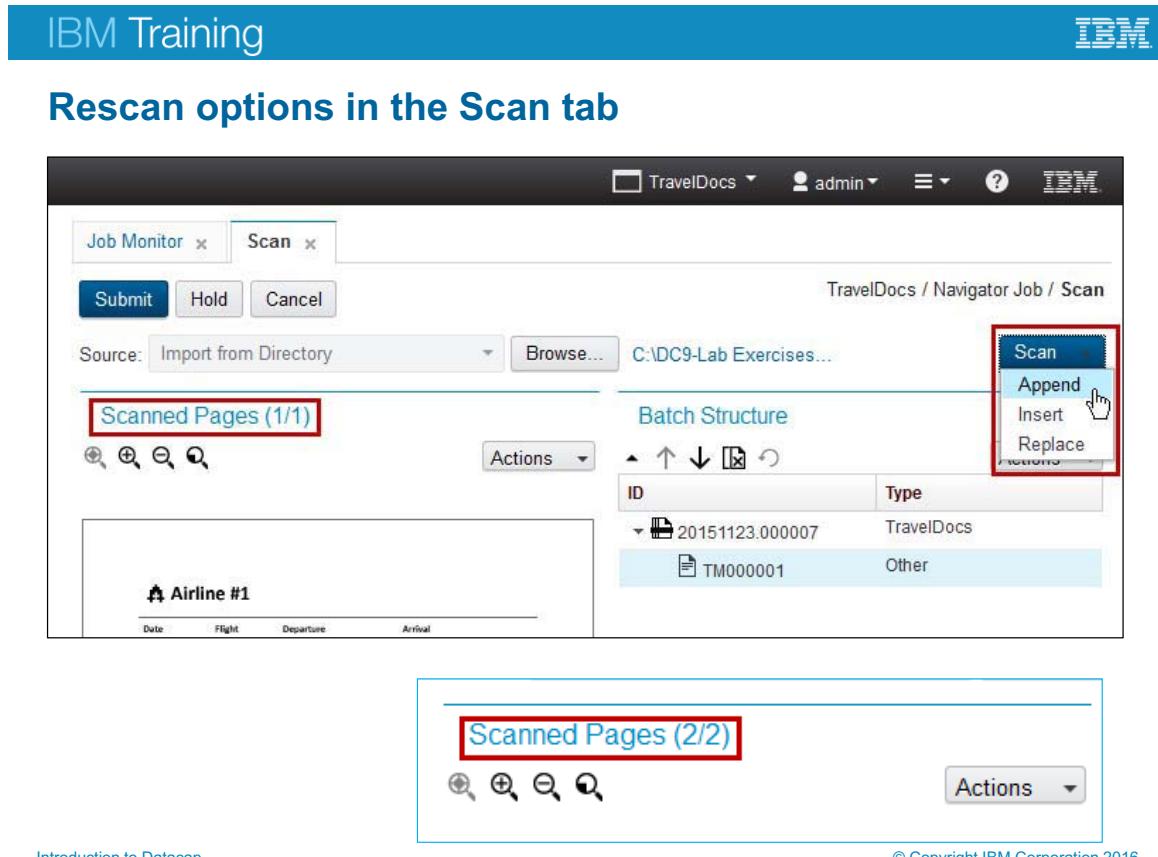


Figure 1-126. Rescan options in the Scan tab

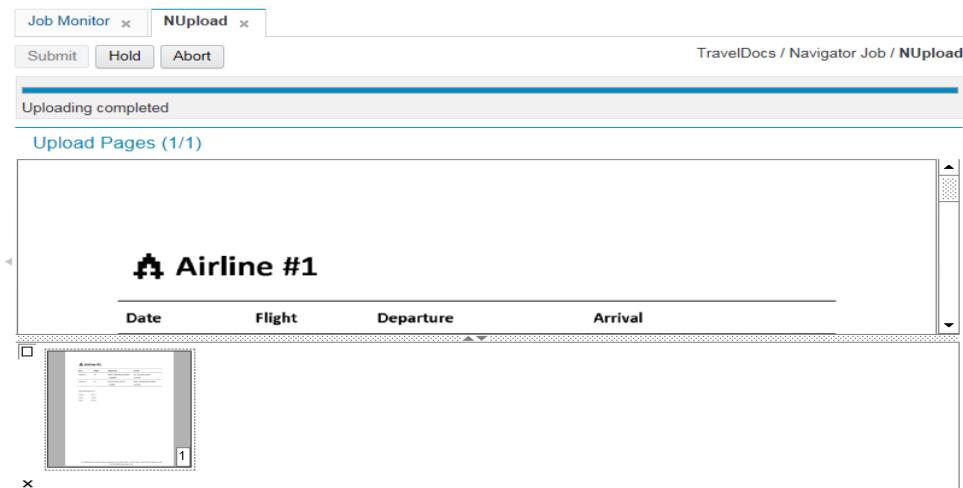
The first screen capture shows the options for rescan.

After you scan the first image, you can do a rescan: "Append", "Insert", or "Replace"

The screen capture at the end of the page shows that after an image is appended, "Scanned Pages" number changed from 1 to 2.

Upload Batches, Page ID, and Profiler tasks

- After you scan a batch in Datacap Navigator, the Upload task is run automatically.



- Page ID and Profiler tasks also run automatically.

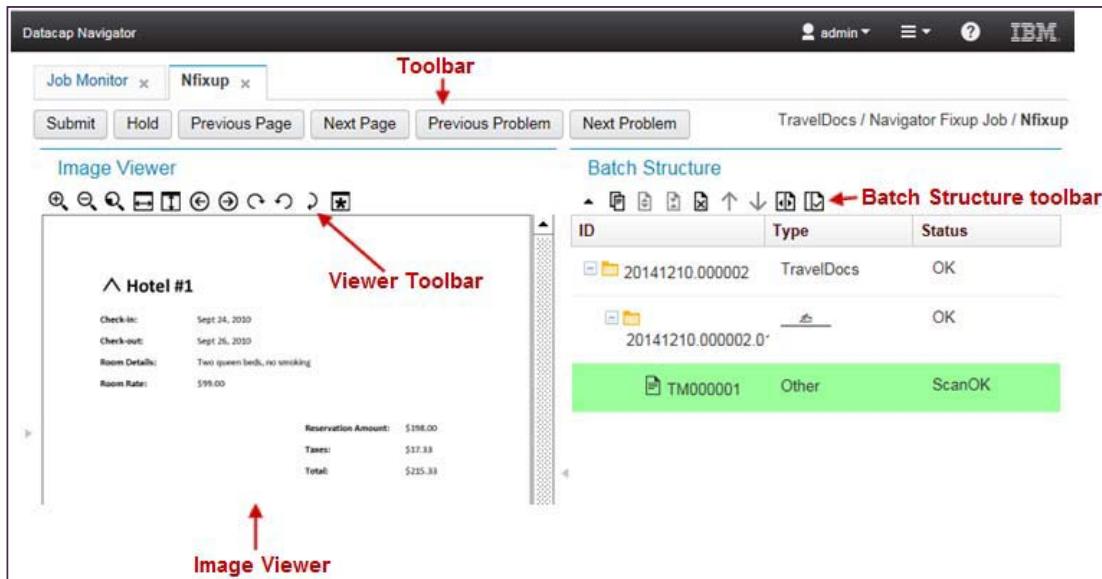
Figure 1-127. Upload Batches, Page ID, and Profiler tasks

You can also run the Upload task manually after scanning a batch in Datacap Navigator.

IBM Training



Classify Batches



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-128. Classify Batches

Help path

- Datacap > Datacap 9.0.1 > IBM Datacap V9.0.1 documentation > Accessibility features of Datacap Navigator
- Toolbar
 - Submit Batch, Hold Batch, Previous/Next Page, Previous/Next Problem
- Batch structure toolbar
 - Drag pages to move across documents
 - Expand, Copy page, Split/Merge documents, Move up/down, Disassemble documents, Check integrity
- Image Viewer toolbar
 - Zoom In/Out/Quarter, Fit to Width/Height, Previous/Next Page, Rotate Clockwise/Counterclockwise/180 Degrees, Show viewer in dual monitor
- Full hot key support. For more details, see Help Path reference.

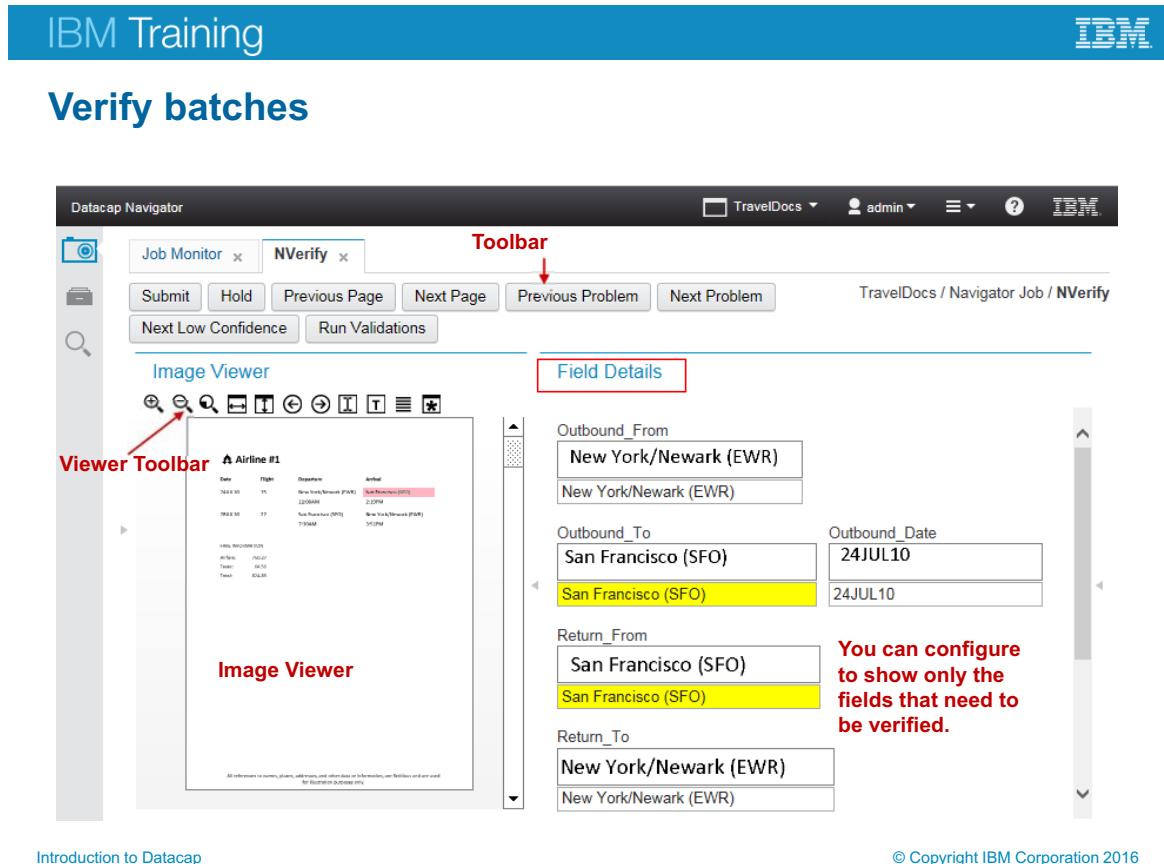


Figure 1-129. Verify batches

Help path

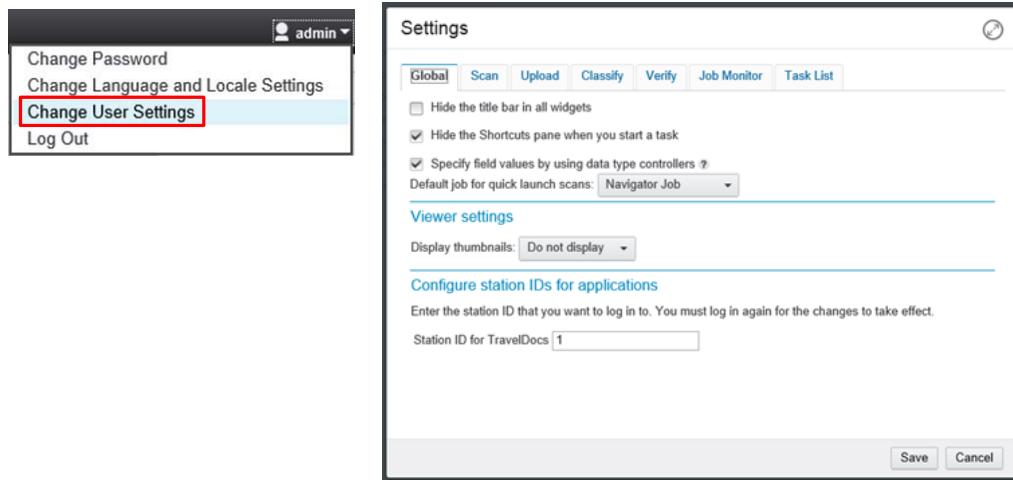
- Datacap > Datacap 9.0.1 > IBM Datacap V9.0.1 documentation > Accessibility features of Datacap Navigator
- Toolbar
 - Submit batch, Hold batch, Previous/Next Page, Previous/Next Problem, Next Low Confidence Field, Run Validation on current page
- Image Viewer toolbar
 - Highlight the area of the field that is selected on in Field Details panel.
 - Update the value for a field with clicking or drawing a rectangle on some words in the viewer. This method of identifying a field is also referred to as the click'n'key method.
 - Zoom In/Out/Quarter, Fit to Width/Height, Previous/Next Page, Show all fields, Show recognized words, Show recognized lines, Show viewer in dual monitor
- Field Details
 - Show both image snippet and value of the field
 - Yellow background: Low confidence fields

- Light pink background: fields that fail validation
- Browser side data type validation
- Support to look up data from database
- OMR – check box and radio buttons
- Batch structure toolbar
 - Expand, Split/Merge documents, Delete page, Move up/down, Disassemble documents, Check integrity
- Full hot key support. For more details, see Knowledge Center Help Path above.



User settings

- Start the User settings from the top bar.
- Each user can change the settings.
 - To tailor the experience to meet the individual needs.
 - To change the appearance and operation of the user interface.



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-130. User settings

Some of the common options that you can configure in User Settings are:

- Hide shortcuts pane when you start a task.
- Show or hide thumbnails for preview
- Set the number of pages to scan for a batch
- What columns are available in the Job Monitor

Accessibility features of Datacap Navigator interface

- Datacap Navigator includes accessibility features.
- Features available for keyboard input and navigation:
 - Keyboard input
 - Keyboard focus
 - Keyboard navigation

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-131. Accessibility features of Datacap Navigator interface

Help path

- Datacap > Datacap 9.0.1 > IBM Datacap V9.0.1 documentation > Accessibility features of Datacap Navigator
- Datacap > Datacap 9.0.1 > Installing > Installing and configuring in a client/server environment > Datacap installation command-line parameters > Commonly used Datacap Setup.exe parameters
http://www.ibm.com/support/knowledgecenter/api/content/nl/en-us/SSZRWV_9.0.1/com.ibm.dc.install.doc/deref003.htm

Keyboard input:

You can use the keyboard instead of a mouse to operate Datacap Navigator.

- To use any icon on the user interface, move to the icon and press the Enter key.
- To enter data:
 - Move to the input field, enter data, and press Enter or press the Tab key to exit the field.

Keyboard focus

The position of the keyboard focus is outlined or highlighted, indicating which area of the window is active and where your keystrokes affects.

Keyboard navigation

You can use the Tab key, Shift+Tab, and Arrow keys to move around the major elements of a page, view, or specialized section.

For certain elements, such as tree views or the calendar date picker, you can also use the Home, End, Page Up, and Page Down keys.

Exercise: Explore the Datacap Navigator interface and process a batch

Requires:
Course Exercise Guide
Student system

Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-132. Exercise: Explore the Datacap Navigator interface and process a batch

Exercise objectives

- Explore the Datacap Navigator Interface.
- Process a batch in Datacap Navigator.



Figure 1-133. Exercise objectives

Lesson 1.9. Datacap web client (tmweb)

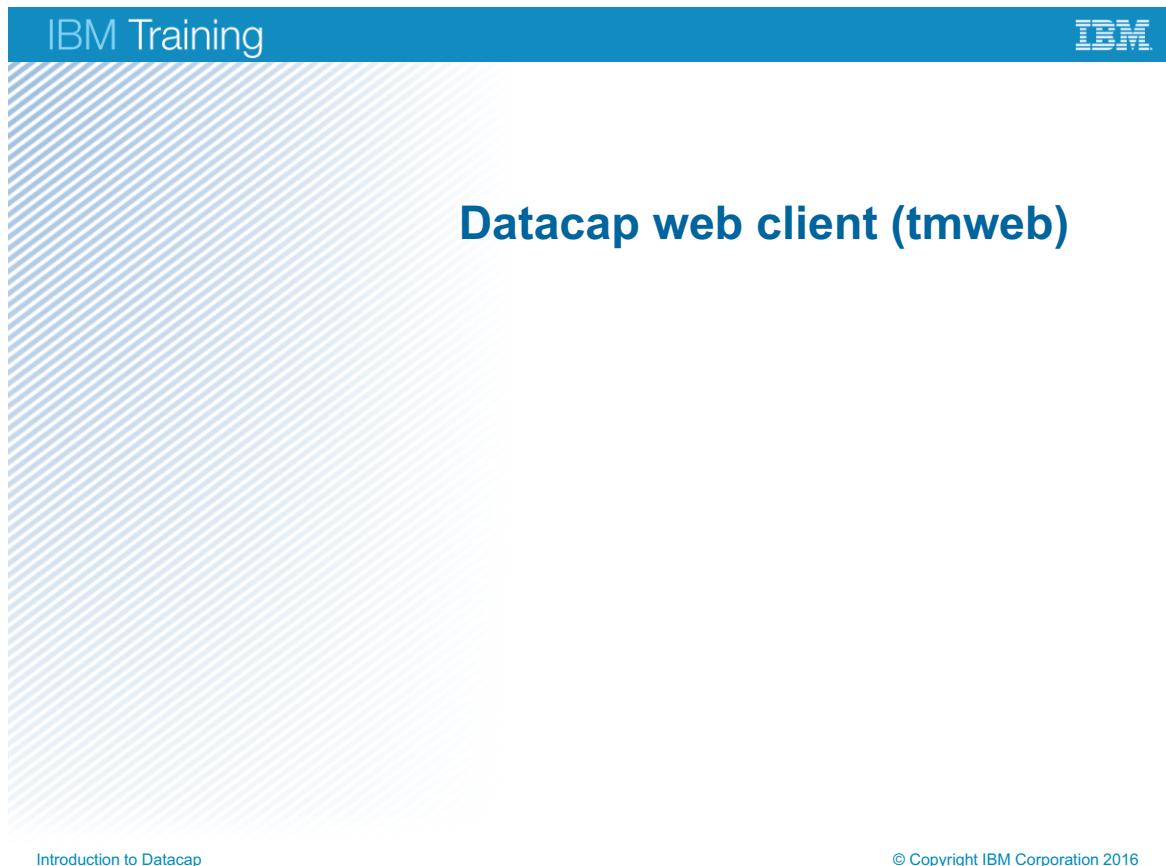


Figure 1-134. Datacap web client (tmweb)

Lessons

- Datacap overview
 - Datacap process
 - Role-based Datacap clients
 - Architecture configurations
 - Architecture components
 - Datacap Desktop
 - Application design
 - Introduction to Datacap Navigator
-  Datacap web client (tmweb)

[Introduction to Datacap](#)

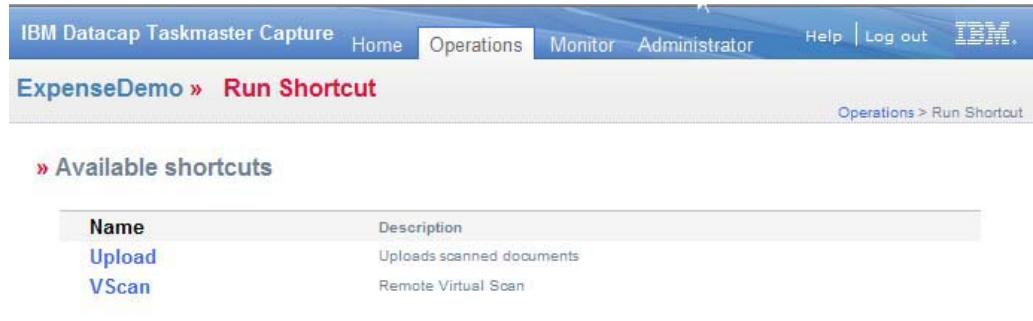
© Copyright IBM Corporation 2016

Figure 1-135. Lessons

Why is this lesson important to you?

- As an administrator of an IBM Datacap capture system, you must be familiar with all configuration tasks for a functional IBM Datacap 9.0 system.
- As an application builder you must know how to use the capabilities of the Datacap Web client. Datacap Web is used for configure workflow jobs and tasks, application users and group, and task shortcuts. Datacap Web can also be used for running capture tasks from remote locations and monitoring active tasks.

Figure 1-136. Why is this lesson important to you?



The screenshot shows the IBM Training Datacap Web Client interface. At the top, there's a blue header bar with the IBM logo and the text "IBM Training". Below the header, the title "Datacap Web Client" is displayed in a large, bold, blue font. Underneath the title, there's a bulleted list of features:

- Operations – Task Processing.
- Monitor – Job and Task progress and statistics.
- Administrator – Configuration

The main content area has a header "IBM Datacap Taskmaster Capture" with tabs for Home, Operations (which is selected), Monitor, Administrator, Help, and Log out. The "Operations" tab has a sub-header "ExpenseDemo » Run Shortcut". Below this, there's a link "Operations > Run Shortcut". The main content area shows a section titled "» Available shortcuts" with a table:

Name	Description
Upload	Uploads scanned documents
VScan	Remote Virtual Scan

Introduction to Datacap © Copyright IBM Corporation 2016

Figure 1-137. Datacap Web Client

Log in to Datacap Web Client

Use Internet Explorer to:

- Browse to <http://localhost/tmweb.net>.
- Log in using a valid user for the authentication mode that you selected.

Operations – Task Processing

The Operations tab displays shortcuts to the tasks in a Datacap application workflow that are configured to run through the Datacap Web client. You can start a task by clicking the name of the shortcut.

Tasks that are configured to run in Datacap Desktop, or by Rulerunner, do not appear on the Operations tab. You must complete these tasks in the program to which they are configured to run.

Monitor – Job and Task progress and statistics.

During the data capture process, documents go through a workflow that consists of several discrete tasks: scanning, upload (if scanned from a remote client), page identification, recognition, validation, verification, and export. Datacap uses a queuing mechanism to move batches of documents through the workflow. On the Job Monitor tab, you view the status of all batches.

To open the Job Monitor, click the Monitor tab in the Datacap Web Client.

Administrator – Configuration

On the Taskmaster Web Client Administrator tab, you configure your application and the application components.

All configuration that is done on the Administrator sub menus is stored in the Admin database. Default is C:\Datacap\<application>\<app name>Adm.mdb.

- Workflow tasks are stored in task table
- Users are stored in tmuser table
- Groups are stored in tmgroup table
- Stations are stored in the station table
- Shortcuts are stored in the Buttons table

Operations – Task Processing

- Use the Datacap Web Operations tab for manual tasks.
- Use Rulerunner for background tasks.
- Factors that determine which tasks are on the Operations menu:
 - Shortcuts that are defined on the Administrator > Shortcuts tab.
 - Tasks that are defined on the Administrator > Workflow tab for web operation.
 - User, Group, and Station permissions.
- Default Operations tasks that the Datacap Application wizard creates are:
 - Upload (Using uplbfcl.aspx)
 - VScan (Using scanc1.aspx)

Figure 1-138. Operations – Task Processing

You can run a batch through the entire workflow by using a combination of web components and Rulerunner.

Web Components can be configured on the operations menu to run tasks that require operator intervention.

Rulerunner is configured to run background tasks that do not require operator intervention.



Monitor – Job and Task Progress and Statistics

- Use the Datacap Web Monitor tab to monitor the status of the job queues.
 - Configure the Monitor tab view:
 - By using the following options.
- Items per page **15** ▾ Delete batches Filter... Refresh rate ▾ Default
- By selecting columns to view.
 - Select several of the column values to filter queued jobs.
 - View job details by clicking the Batch number link
 - Roll back Jobs to a previous task to reprocess tasks.
 - View the task history for the queued batches.

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-139. Monitor – Job and Task Progress and Statistics

Configure the Monitor tab view.

- **Items per page** controls how many jobs are displayed
- **Delete batches** deletes all displayed batches (use the Batch, Job, Task, and Status fields to control which jobs are displayed, or use the Filter link)

Tip: To delete an individual batch, click the batch number and then click Delete.

Tip: The ability to delete batches is security driven so that only privileged users can delete.

- **Filter** provides finer control over which jobs are displayed
- **Refresh** refreshes the job list (or set the rate to refresh automatically)
- **Default** returns to the default view (all jobs)

Column View Links:

- Across the bottom of the screen, is a link for each column of batch data that is viewable. By clicking the column links, you activate or hide the batch data columns.

The following filters filter the Queued Jobs.

You select each filter by clicking Filter... and selecting the option on the top row as shown on the slide screen capture image.

- Batch
- Job
- Task
- Status
- Operator
- Station

View job details by clicking the Batch column link

Roll back Jobs to a previous task to reprocess tasks.

View the task history for the queued batches.

Administrator – Configuration

- Workflow
 - View and configure the application's workflows, jobs, and tasks.
- Groups
 - Configures user groups and allowed tasks for each group.
- Users
 - Configures users and allowed tasks for each user.
- Stations
 - Configures stations and allowed tasks for each station.
- Shortcuts
 - Configures the batch selection mode and the icons that are displayed in the Operations window.
- QA
 - Allows an administrator to create, configure, or delete quality assurance jobs that are based on groups and tasks.

Figure 1-140. Administrator – Configuration

Workflows – Job and Task Definition

- Configuring Jobs
 - Job details:
 - Name (Main Job, Web Job, Fixup Job)
 - Description
 - Priority
- Configuring Tasks
 - Task details:
 - Name (VScan, PageID, Profiler, Verify, Export, Fixup, iVScan, Upload)
 - Description
 - Mode
 - Queue by
 - Store

Figure 1-141. Workflows – Job and Task Definition

Mode:

- Batch Creation: Select this mode for use with VScan or if you are creating a task to scan hardcopy documents from DotScan.
Important: A job can contain only one Batch Creation task. If the job that you are modifying already includes a batch creation task, you must remove that task.
- Router: Select this mode if the task routes the batch to a different task or job when the criteria of a condition are met. One example of a condition is a document integrity failure that requires a supervisor's intervention.
- Normal: This mode is for all other tasks that are not used for Batch Creation or that do not require special handling.

Queue to:

You can specify whether to queue the batch to a stored user, a stored workstation, or both, when it reaches this task in the workflow. If queuing is not a requirement, select None.

Store:

When you select a Store attribute on a task, Taskmaster stores the user or station (or both) that is running the task when a batch is processed. Later in the workflow this stored value is used to

send the batch back to the same user or station (or to a different user or station) in a subsequent task. This control is typically used for one of the following use cases:

- Return documents to the original scan operator (or department) for verification.
- Force a different operator to do a second pass of verification on the document.

Workflows – Task Setup

- Parameters
 - Program Value:
 - For a task run exclusively from the tmweb, select an .aspx page.
 - For a task run from an application client, select the client. Datacap Desktop, FastDoc, Mobile Capture
 - For a task run exclusively in the background, select Rulerunner. PageID, Profiler, and Export are run in Rulerunner.
 - For a task run by Rulerunner, tmweb, or an application client, select Multiple.
- Setup
 - Use Setup to provide more Batch Processing, Rulerunner, Scanning and other options that are related to the task.
 - The parameters that are provided in the setup window depend on the Program Value that is selected.

Figure 1-142. Workflows – Task Setup

Program Value Menu Options

- Application Clients
 - Datacap Desktop
 - FastDoc
 - Mobile Capture
- Background
 - Rulerunner
- Datacap Web Pages
 - Scanl.aspx – Hardcopy remote scan with ISISScan or TWAINScan
 - VScanl.aspx – Virtual Scan of graphic images.
 - UpIBFcl.aspx – Upload batch images.
 - Pickup.aspx – Batch creation task for Datacap Web that involves different types of documents, such as graphics, bar codes, text, and forms.
 - ProtolD.aspx – Manual Page Identification and fixup.

- Restruct.aspx Manual identification of page type and reassembling documents.
- aIndex.aspx – Verification, manual page identification, and manual registration
- aVerify.aspx - Verify
- ImgEnter.aspx - Verify
- VeriFine.aspx - Verify

Add Users and Groups

- Adding users to an application.
 - Add Datacap users for TMA.
 - Password is only required with TMA authentication
 - To add users, you need User privilege.
- Adding groups to an application.
 - Add Datacap groups for TMA authentication.
 - To add groups, you need User groups privileges.
- Authentication systems that use LDAP and Active Directory are covered in the Datacap 9.0 System Configuration class.

Figure 1-143. Add Users and Groups

The internal authentication system is Task Master Authentication (TMA).

Adding users to an application

Defining users enables users to work on the Datacap system. They must have a user ID to authenticate with. The user-defined privileges, permissions, and group associations determine what the user is allowed to do.

- TMA is the only authentication system that uses the Datacap user for authentication. Therefore, it is the only system that requires a user for the Datacap user.

Adding groups to an application

Defining groups enables groups of users to be associated and to all have the same access credentials.

Add Users to Groups and Add Stations

- Add users to groups.
 - To create a group, you need User groups privilege.
 - To add a user to a group, you need User groups privilege and you must be a member of the group.
- Add stations to an application
 - Stations are not necessarily tied to a specific system.
 - Stations can be configured to allow multiple users to log in on one station.
 - Set Maximum to control the maximum number of users that are allowed.
 - Station configuration allows Permissions configuration.

Figure 1-144. Add Users to Groups and Add Stations

Adding users to groups

Any user can create a group provided they have the User groups privilege. When the group is created; only a user who is a member of the group and also has User groups privilege can modify it. The restriction of allowing only the member of a group to change a group is by design. There is a work-around method for an administrator who is not a group member. You can select the option to copy a group, and make your wanted changes. Then, delete the old group and rename the new group to have the original name.

Adding stations to an application

The Stations tab within Datacap Web provides you with the ability to create a station, and assign it a unique identifier. You can also define which applications, workflows, and job-task pairs can be run when a user logs in to Datacap with that station ID.

A station ID does not have a one-to-one correspondence with a physical workstation. You can enable the Datacap feature called virtual stations. Virtual stations are set up by setting the Maximum number greater than zero. When a virtual station is configured then different users them to log in on different physical workstations simultaneously.

When a virtual station is configured, multiple users can log in on different physical workstations with the same virtual station definition, simultaneously. When the Maximum number of virtual stations is set to zero, Datacap prevents multiple users from logging in with the same station ID.

Datacap Web sessions can timeout, and users can close their browser windows without logging out properly. When these occurrences happen, setting a Maximum number of virtual stations that are greater than zero allows users to log back in. Otherwise, system administrator support is requiring to log back in. When the Maximum number of virtual stations for a station ID is set to zero, the user's next attempt to log in fails. Also, if the Maximum number of virtual stations is reached the next login fails. In this situation, the system administrator must clear the virtual stations for that station ID to allow the user to log in again.

Set Privileges

- Privileges determine the valid user actions.
- Privileges are set at group or user level and are cumulative.
 - Total privilege = Group privilege + User privilege.
- Privileges are grouped into sets.
 - Job Monitor
 - Administrator
 - Station/Web monitor
 - ‘Run Task’ dialog
 - General settings dialog
 - Communications
 - Clients

Figure 1-145. Set Privileges

Set Privileges

Selecting privileges determines specific actions that the user or group that you are configuring can do. Privileges are arranged in sets.

The following are privilege sets:

Job Monitor

- This set configures the job monitoring actions a user can control.

Administrator

- This setting determines the system administrative functions a user can do.

Station/Web monitor

- This privilege allows access to the Station monitor and Web monitor view in the Datacap Web client.

‘Run Task’ dialog

- This setting allows access to the Report Viewer, Datacap Web, and Datacap Studio clients.

Set Permissions

- Permissions define what job tasks the user can run.
- Permissions can be selected for users, groups, and station configuration.
- Task permissions are grouped by job on user, group, and station property pages.
 - Main Job – Defines when tasks can run with DotScan and Rulerunner
 - Fixup Job – Defines when the Fixup job can run.
 - Web Job – Defines when a Web Client can run Jobs.

Figure 1-146. Set Permissions

Configure Shortcuts

- Shortcut Details
 - Name
 - Description
 - Mode
 - Prompt/Web select
 - Auto
 - Manual
 - Manual for Hold
- Permissions
 - Permissions selections map each Shortcut to tasks.

[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-147. Configure Shortcuts

Mode descriptions

For the Mode, select one of these values that determines the behavior of Datacap when a user clicks the shortcut on the Workflow tab:

- **Prompt/Web select:** Datacap opens the highest priority-pending job in the queue.
Tip: You can change the priority of a batch on the tmweb > Monitor view to control the order of batch processing.
- **Auto:** Same as Prompt/Web select.
- **Manual:** Datacap displays the job queue so that the operator can select a batch, which is either pending or on hold.
- **Manual for Hold:** If there are batches on hold, Datacap displays the job queue with the jobs that are on hold. If there are no jobs on hold, Datacap opens the highest priority-pending batch in the queue.

Review questions (1)

1. True or False.
tmweb is the Datacap Capture application development environment.
2. True or False
tmweb is one of the capture job processing environments.
3. True or False
tmweb provides direct access to the Application Wizard through an icon on the icon bar.
4. True or False
tmweb provides the capability to process document batches manually through every task of the capture process.

Figure 1-148. Review questions (1)

Review answers (1)

1. True or False.
tmweb is the Datacap Capture application development environment.
[The Answer is: False](#)
2. True or False
tmweb is one of the capture job processing environments.
[The Answer is: True](#)
3. True or False
tmweb provides direct access to the Application Wizard through an icon on the icon bar.
[The Answer is: False](#)
4. True or False
tmweb provides the capability to process document batches manually through every task of the capture process.
[The Answer is: False](#)

Figure 1-149. Review answers (1)

Review questions (2)

5. True or False

tmweb Client provides the capability to process document batches manually through some task of the capture process.

6. True or False

Taskmaster Administrator functions are done on the tmweb > Administrator tab subtabs.

7. True or False

Taskmaster Administrator functions are done on the tmweb > Configuration tab subtabs.

8. True or False

The tasks on the tmweb > Operations tab are configured on the tmweb > Administrator > Workflow tab and the tmweb > Administrator > Shortcut tabs.

Figure 1-150. Review questions (2)

Review answers (2)

5. True or False

tmweb Client provides the capability to process document batches manually through some task of the capture process.

The Answer is: [True](#)

6. True or False

Taskmaster Administrator functions are done on the tmweb > Administrator tab subtabs.

The Answer is: [True](#)

7. True or False

Taskmaster Administrator functions are done on the tmweb > Configuration tab subtabs.

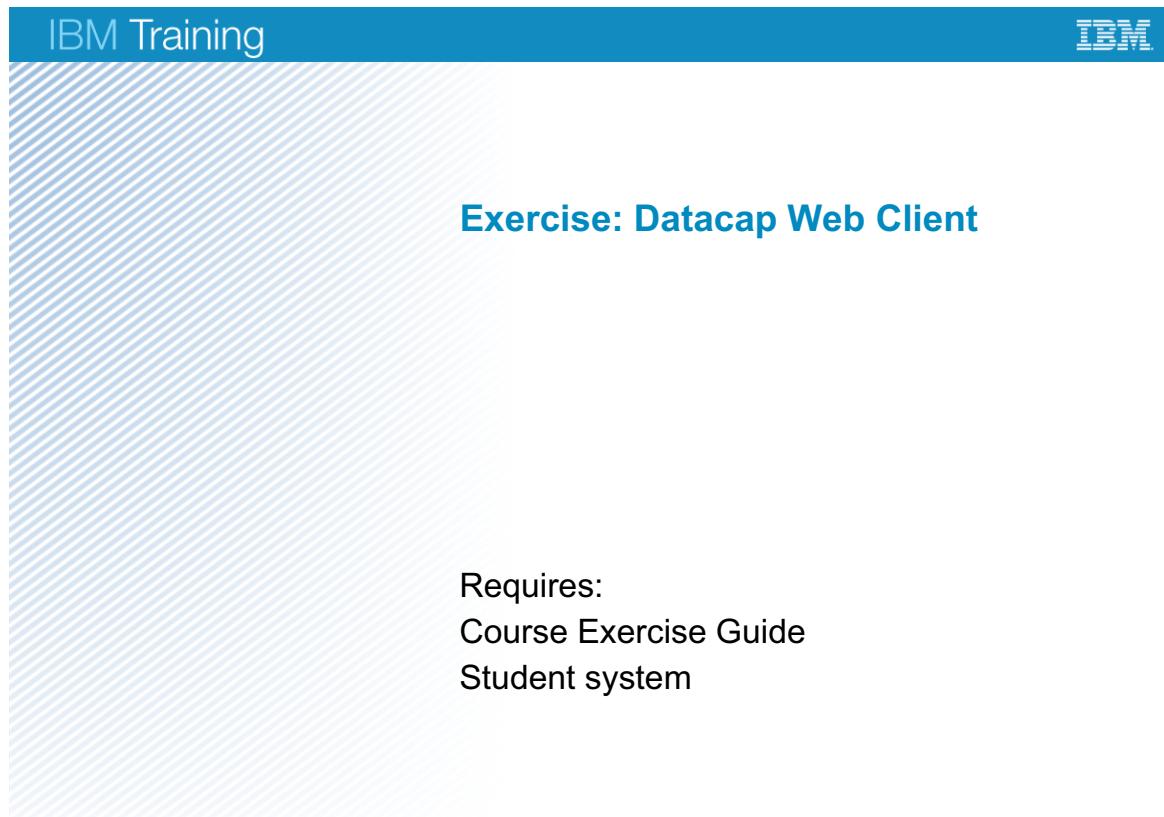
The Answer is: [False](#)

8. True or False

The tasks on the tmweb > Operations tab are configured on the tmweb > Administrator > Workflow tab and the tmweb > Administrator > Shortcut tabs.

The Answer is: [True](#)

Figure 1-151. Review answers (2)



Introduction to Datacap

© Copyright IBM Corporation 2016

Figure 1-152. Exercise: Datacap Web Client

Exercise objectives

- Explore the Datacap Web Client interface



[Introduction to Datacap](#)

© Copyright IBM Corporation 2016

Figure 1-153. Exercise objectives

Unit summary

- Identify the business solution that IBM Datacap provides, the Datacap process, and the capabilities of Datacap.
- Work with the Datacap Navigator, Datacap Desktop (Windows based client), and the Datacap web client (tmweb) to process a batch of input data.
- Identify the components of Datacap (Architecture) and the things to consider for Application Design.

Figure 1-154. Unit summary

Unit 2. FastDoc Local Mode

Estimated time

04:00 Hrs

Overview

This unit introduces you to the Datacap Fastdoc. It then proceeds to focus on the Local mode batch and task configuration and local and upload processing.

How you will check your progress

- Successfully complete the activities in the Student Exercises book

References

IBM Knowledge Center

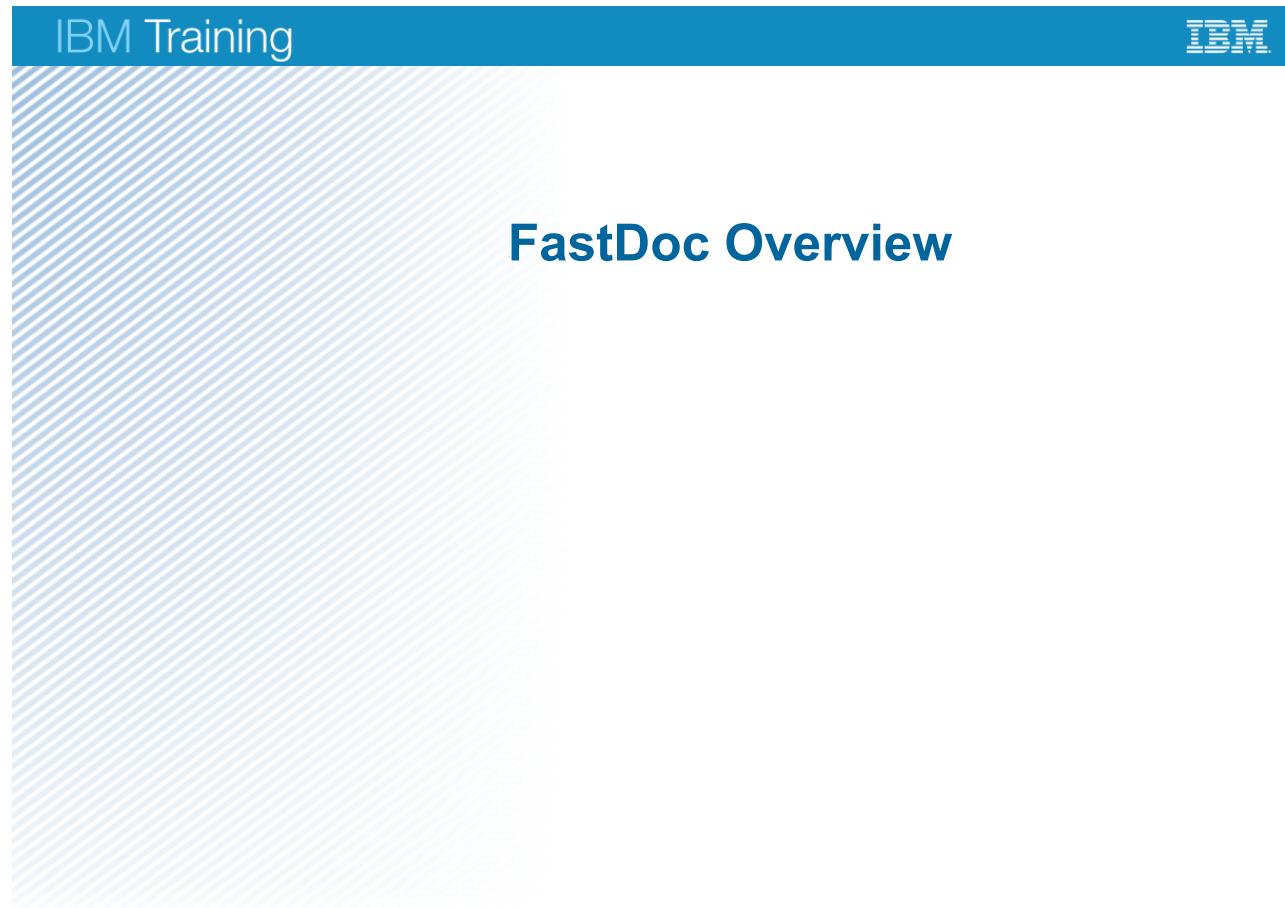
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/datacap_9.0.1.htm

Unit Objectives

After completing this unit, you should be able to:

- Identify the operating modes of Datacap FastDoc.
- Configure batches for off line processing.
- Scan and upload a batch

Lesson 2.1. FastDoc Overview



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-2. FastDoc Overview

Lessons

FastDoc Overview

- FastDoc local mode
- Create a local capture batch
- Process local batches
- Configure Scan and Upload

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-3. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system.
- This lesson gives you a broad overview of the capabilities of the FastDoc as a rapid development tool and as a user client.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-4. Why is this lesson important to you?

What Is FastDoc?

- It is a flexible Datacap interface with three primary functions:
 - Rapid application development with the FastDoc (Admin) interface.
 - Process batches partially or completely in the Local - Offline mode on the local machine.
 - Process batches with the online Datacap server mode as a client.
- FastDoc operation modes
 - Local – Offline Mode
 - Datacap Server – Online Mode
 - FastDoc is used as a client interface for manually running a task for a batch that is in the Datacap Server queue.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-5. What Is FastDoc?

FastDoc operation modes

FastDoc has two modes of operation:

- Local - Offline Mode
 - Used for processing batches partially or completely, on the local machine before uploading to Datacap server.
 - The Local mode can be configured to run batches for multiple applications.
- Datacap server - Online Mode.
 - In the Datacap server mode, FastDoc is used as a client interface for manually running a task for a batch that is in the Datacap Server queue.
 - When logging in the Datacap server mode, you are connected to a specific application.

FastDoc Versatility

- Offline development
 - Define offline workflows and tasks.
- Online development
 - With Admin interface, FastDoc functions as a rapid development tool.
 - Configure document, page, and field objects, fingerprints, field location, validation rules, and export rules
- Local – Offline Client
 - Functions as a stand-alone processing or application development tool
- Datacap server – Online Client
 - Functions as client interface to an application on the server or a rapid development tool by using predefined application templates.

FastDoc Local Mode

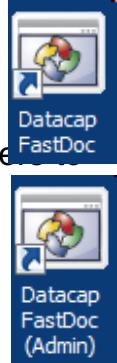
© Copyright IBM Corporation 2016

Figure 2-6. FastDoc Versatility

- FastDoc is most likely to be run in the local mode when it is installed in a workstation that is at a site remote from the Datacap File Server. In this case, a copy of the applications that it runs must be installed at the remote site together with FastDoc.
- FastDoc is run in the Developer Workstations for Business Analysts or Application developers to Create or Configure Applications.
- If FastDoc is intended to run in the Datacap Server, then it is installed on a Client Workstation that is at the site where the Datacap File Server is located. In this case, the application that is built to process the batches does not need to be installed on the Client Workstation. The Client Workstation has access to the Datacap file server where the central application is installed.

FastDoc interfaces

- There are two FastDoc interfaces:
 - User Version is used for processing document in either Local or Datacap server mode.
 - Administrator Version is used by Administrators or Application Builders to configure or process document batches.
- Each interface can run in two modes:
 - Select the run mode from the login window.
- The run modes are:
 - Local mode when:
 - Connection to the server isn't stable.
 - FastDoc is installed at a remote site.
 - Database tracking and sharing application resources are Local Datacap Server
 - Datacap Server mode
 - Run FastDoc as a Datacap Client while for manually running a task for a workflow that is in the Datacap Server queue.



Welcome to Datacap

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-7. FastDoc interfaces

FastDoc interfaces

FastDoc has two operation interfaces:

- User Version, FastDoc
 - For users to process document batches in either Local or Datacap Server mode.
 - Click Start > All Programs > IBM Datacap Clients > FastDoc
- Administrator Version, FastDoc (Admin)
 - For Administrators or Application Builders to configure workflow processing for the local mode.
 - For Administrators or Application Builders to process document workflows in either local or Datacap Server mode.
 - Click Start > All Programs > IBM Datacap Clients > FastDoc (Admin)
- Select the Local or Datacap Server mode on the login window.

FastDoc Interface Icons

FastDoc	Purpose	Explanation
1 	Provides a list of icons for processing documents.	The user can click these shortcuts to process documents.
2 	Configure documents, pages, and fields Admin mode only	This option is used to design and configure documents, pages, and fields.
3 	Configure workflow Admin mode only	This option is used to design and configure jobs and tasks.
4 	Configure local batches Admin mode only	This option is used for configuring local batches for offline processing.
5 	Logout	To logout from FastDoc
6 	Application Wizard	For Creating New Applications

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-8. FastDoc Interface Icons

- With the User interface in either the local or Datacap Server mode, you see only icon 5. You are always presented with the batch processing view.
 - In the local mode, you can run only local workflows that are already configured.
 - In User Datacap Server mode, you run only task from the Shortcuts list.
- In Admin local mode, you can see 1, 4, 5, and 6.
 - Use 1 to Process batches.
 - Use 4 to configure local workflows.
- In Admin Datacap Server mode, you can see 1, 2, 3, 5, and 6.
 - Use 1 to Process batches on the server.
 - Use 2 and 3 to configure the batch structure, fingerprints, and rulesets. (Build the Application)



FastDoc overview (1)

- Before you begin:
 - Decide whether you want to use the Admin or the User interface.
 - The same login window is used for both interfaces.
 - The login Credentials are only required for Datacap Server mode.

Welcome to Datacap

Local Datacap Server

Log in credentials not required.

Login

IBM.

Welcome to Datacap

Local Datacap Server

Application

FastBuild

User

admin

Password

Station

1

IBM.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-9. FastDoc overview (1)

Images on the slide show the FastDoc login window.

When to use?

- You are always prompted to log in every time you start one of the FastDoc interfaces.
- If you select the Local option, the credential fields go away. In the Local mode, you select the batch that you want to configure or run inside the interface.
- If you select the Datacap Server mode, then you are required to select the specific application to connect to and to provide valid credentials for the application.

Before you Begin

- Decide whether you want to use the Admin or the User interface.
- Decide whether you want to run in the Local - Offline mode or the Datacap Server mode.

FastDoc Local login

- For either (Admin) or (User) interface, you can select the Local - Offline mode (no login credentials are required).
 - Click Local and click Login
- In the Local mode, you can run a batch of any application that has a defined batch profile.

FastDoc Datacap Server login

- For either (Admin) or (User) interface, you can log in to the Datacap Server - Online mode by selecting an application, providing login credentials, and a station filter and then clicking Login.

Application: (Select application from list)

User: admin

Password: admin

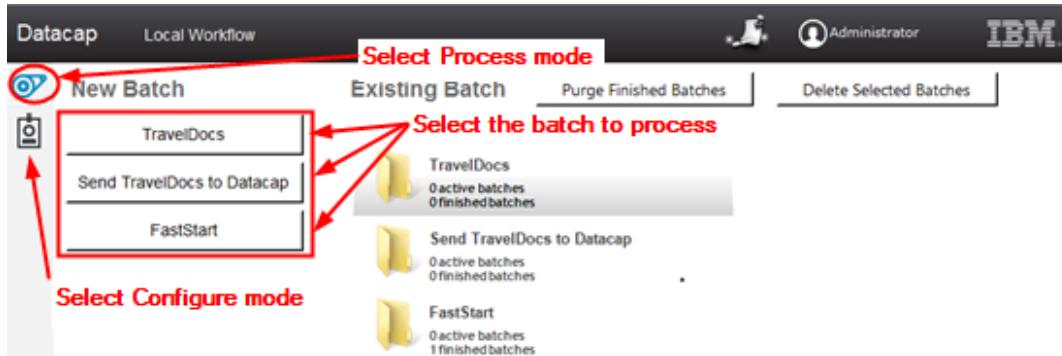
Station ID: 1

- Click Login.

- In the Datacap Server mode, you log in to a specific application based on station filter description and configured privileges and permissions.
- You can run only tasks for the application you are logged in to.



FastDoc overview (2)



- Batches are defined in the Configure mode.
- The Configure mode is only available in the Admin interface.
- Batches automatically progress from one task to the next until the workflow is completed.
 - Only one Batch link is required to run all of the tasks.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-10. FastDoc overview (2)

The image on the slide shows the Local mode Process batches view.

- Only one link is required to run all of the tasks in a batch because when the batch is initiated FastDoc automatically runs all of the tasks that are defined for the batch.
- There might be multiple batches that are defined for each application that processes batches in the local mode. For example:
 - One batch might be configured to scan only and then upload the batch to the Datacap File Server system for further processing.
 - One batch might be configured to do all processing locally, scan, indexing, validation, and verification.

FastDoc overview (3)

- Can use the same interface for:
 - Local offline processing and upload.
 - Datacap server mode online processing after upload.
- Can be used as an application configuration tool or client for processing batches.
 - Can switch between development mode and batch processing mode without switching interfaces.
- With the help of built-in templates, you can quickly create and configured applications.
- Can create and save your own custom templates.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-11. FastDoc overview (3)

Use Datacap Desktop or Datacap Navigator if you require the following functions:

- Custom Verify panels.
- Line item verification.
- SELECT and Lookup DCO properties for database queries.

FastDoc Terms (1)

FastDoc (Admin)	FastDoc administrative interface. Start>All Programs>Datacap Development Tools>Datacap FastDoc (Admin)
FastDoc	FastDoc user interface. Start>All Programs>Datacap Clients>Datacap FastDoc
Local mode	FastDoc mode for configuring and running local or offline batches.
Datacap Server mode	FastDoc mode for configuring and running Datacap workflows as a Datacap client.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-12. FastDoc Terms (1)

FastDoc (Admin)

The administrative interface is for configuring tasks and batches for processing documents batches.

- In the local mode, local batches and tasks are configured.
- In the Datacap Server mode, complete applications are created and configured by using a template or another application as a starting point.

FastDoc – Local mode

All processing is done automatically except where human intervention is required.

Operator intervention is required at the Verify step.

After validation completes at the Verify task, the batch advances through the Export task.

FastDoc – Datacap Server mode

Each defined batch task must be manually initiated.

Scan, Profiler, Verify, and Export.

Background tasks like Profiler and Export can be configured to run automatically with the Rulerunner service.

FastDoc Terms (2)

Batch profiles	An offline workflow that runs in FastDoc Local mode to process a batch through all tasks. When initiated, it processes all tasks in sequence and stops only when operator intervention is required.
Job	A Datacap Server workflow that consists of a number of Tasks.
Task	One step in a job or batch profile.
Task Profile	A sequence of Rulesets that are processed as Mapped to a task.
Batch	A group of documents that are processed together. Also, all of the files that are created in the course of all tasks in a job.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-13. FastDoc Terms (2)

Batch profiles

- In the context of FastDoc running in the Local - Offline mode, processing is done in batches rather than task at a time.
 - When you start a batch, FastDoc automatically walks through each task that is defined in the batch until some operator intervention is required.
 - A batch profile is a collection of tasks that are arranged in the order in which the tasks are run.
 - A batch profile usually starts with a Scan task and ends with the task that is configured with the Send to Taskmaster option.
 - Send to Taskmaster option is explained in a later section.

Job

- The tasks that are associated with a job represent all the processing that must be done on a batch of documents to take the batch through the:
 - Source capture phase.
 - Transformation phase.

- Delivery phase.

Task

Examples of tasks might be: Scan, PageID, Profiler, Verify, and Export.

Task Profile

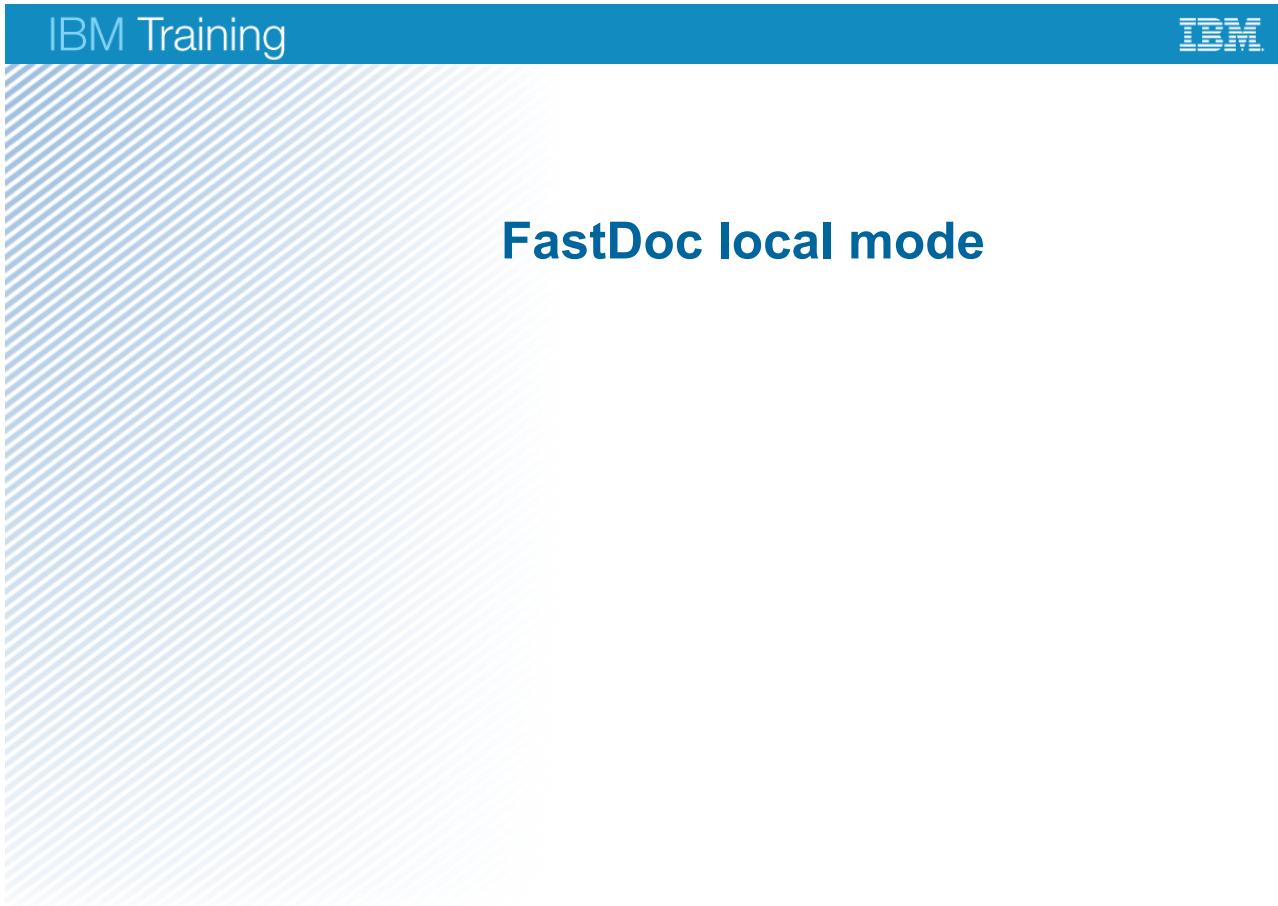
- From the prerequisite Datacap Introduction class, you know that a task profile defines a list of references to rulesets run in a task. Defines what is done on a document batch for each processing step.
- When FastDoc runs in the Datacap Server mode, you run one task at a time.
- Task Profiles are assigned to tasks to define the task process.

Batch

A Batch is a group of pages that are scanned and processed through each Datacap task as an entity. When the entire batch process is complete the result is:

- Pages are assembled into documents and exported to a data repository.
- There is a batch folder on the Datacap Server machine that preserves the history of the capture process.

Lesson 2.2. FastDoc local mode



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-14. FastDoc local mode

Lessons

- FastDoc Overview
-  FastDoc local mode
- Create a local capture batch
- Process local batches
- Configure Scan and Upload

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-15. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you might be required to configure FastDoc to process batches completely or partially on the local system before uploading the batches to the server for further processing.
- To do these tasks effectively, you must be familiar with the FastDoc Local mode interface.

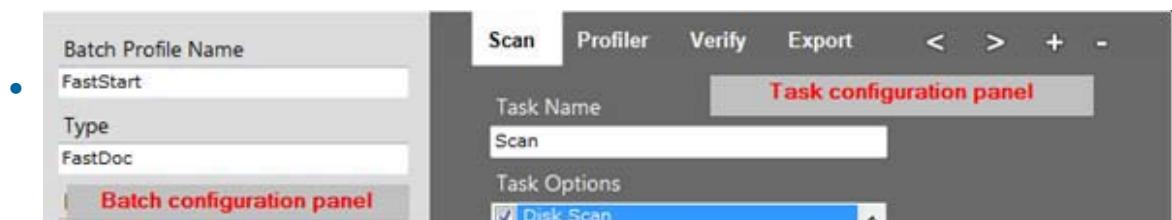
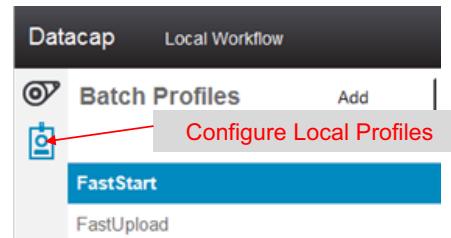
FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-16. Why is this lesson important to you?

FastDoc Local Batch Configuration

- Batch Profiles are only needed in the FastDoc Local mode.
- Batch Profiles are equivalent to Jobs in Datacap Server terms.
- Create local batches in the FastDoc (admin) Local mode.
 - On the Batch Profiles sidebar, click Configure local profile.
- Configure the batch parameters.
- Create a task for each step that you process in the Local mode.



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-17. FastDoc Local Batch Configuration

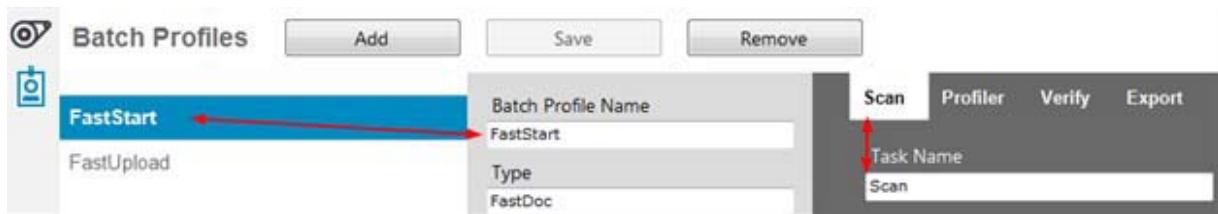
Configuring applications in Local mode

- Start FastDoc (Admin)
- Click Configure local profile
- Click Local
- Click Login
- To Add a new batch click Add and define the new batch parameter.
- To edit an existing batch, click the batch and edit the parameters directly.
- Click Save to preserve your changes when editing is complete.

IBM Training



Configure Local Batch Profiles



- **Batch toolbar**
 - Batch shortcut name follows the Batch Profile Name
 - Use Add to create another batch.
 - Use Remove to delete a batch.
 - Use Save to preserve configuration changes

FastDoc Local Mode

© Copyright IBM Corporation 2016

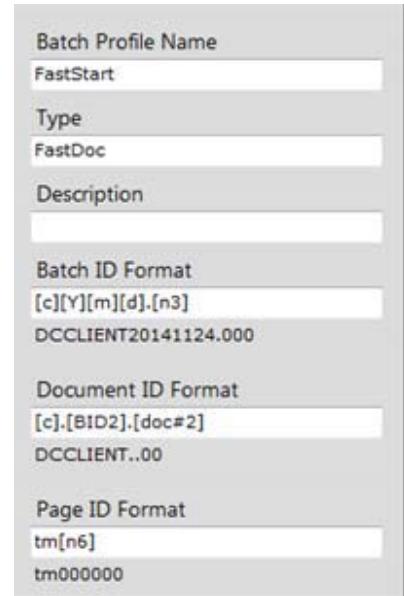
Figure 2-18. Configure Local Batch Profiles

Help path.

- Datacap > Datacap 9.0.1 > Customizing and running Datacap applications > Preparing and running FastDoc applications > Application configuration on FastDoc > Setting up batch profiles in Local mode

Configure the Batch Parameters

- Batch Parameter pane
 - Batch Profile Name.
 - Type
 - Description
 - Batch ID Format
 - Document ID format
 - Page ID Format
 - Format descriptions:
 - c – Computer name
 - Y – four-digit year
 - m – two-digit month
 - d – two-digit day of month
 - n3 – three-digit batch counter
 - BID2 – Batch ID
 - tm fixed text
 - doc#2 – two-digit document number in the batch
 - n6 – six-digit image count within the batch



[FastDoc Local Mode](#)

© Copyright IBM Corporation 2016

Figure 2-19. Configure the Batch Parameters

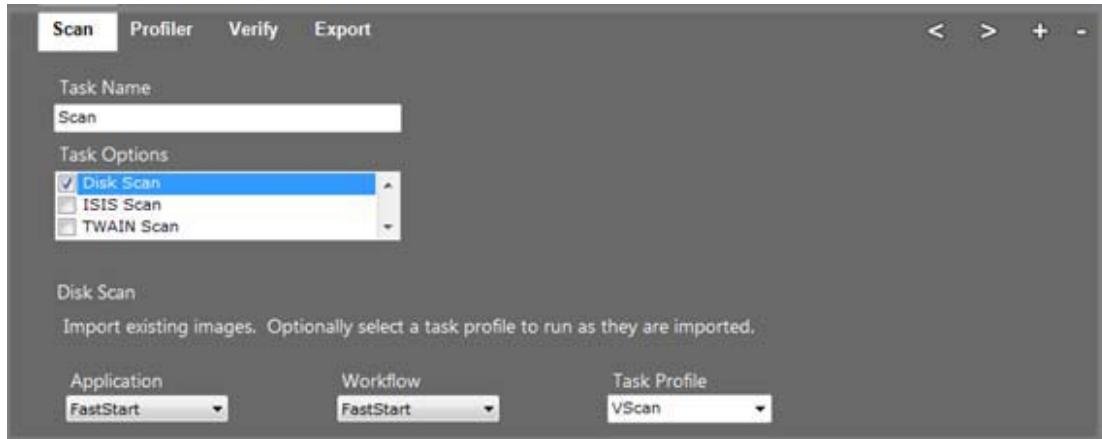
- The entries on this pane determine the parameters for a Batch Profile:
 - The Shortcut name (Batch Profile Name) that starts the execution of the batch in the FastDoc Local mode.
 - The name (Batch ID) and location (Batch Profile Name) of the batches folder.
 - The name that is used for task-related files in the batches folder. (Document ID)
 - The name that is used for image files in the batches folder. (Page ID)
- Batch Profile Name
 - Used for the batch folder name in the FastDoc > batches folder.
C:\Datacap\FastDoc\<Batch Profile Name>\batches\<date>.<num>
 - Used for the Batch Shortcut name that is used to start the batch in the Local mode interface.
 - Coincidentally it is the same name as the Datacap Application name.
- Type
 - Identifies the Document Hierarchy (DCO) that this batch is associated with.
 - This field must match the DCO type or Batch structure name.

- Description.
 - Any descriptive definition.
- Batch ID Format- rules for the folder names
 - Defines the batch folder name.
 - Batch folder is where batch history is accumulated as tasks are processed.
- Document ID Format - rules for document names
 - Document number to append to the BID if there are multiple documents in the batch.
- Page ID Format - rules for page names in the batch



Configuration Local Tasks

- Provide a task name
- Task Options
- Configure more parameters



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-20. Configuration Local Tasks

Help path

- Datacap > Datacap 9.0.1 > Developing applications > Custom Datacap Report Viewer reports > Datacap statistics
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.dc.customrv2.doc/dcapi000.htm?lang=en

Configuring Local Tasks

- Task Name is displayed on the task tab. There is a tab for each Task defined.
- Task toolbar options
 - < > Left and Right. Use these controls to order the task menu tabs.
 - + Plus sign to add a task.
 - - Minus sign to remove a task.
- Task options are:
 - Disk Scan
 - ISIS Scan

- TWAIN Scan
- Send to Datacap
- Verify
- Rulerunner
- Start Batch Panel
- Task Option parameters.
 - Each task option that you can select has a unique set of parameters that you provide.

Examples of required parameters are:

 - Application name
 - Workflow name
 - Task Profile name
 - Job name, task
 - Batch Folder
 - Statistics Table. See the IBM Knowledge Center help link.

Two FastDoc workstation configurations

- On the same LAN as the Datacap Server.
 - Application Manager > Service* = \\<Server>\Datacap\datacap.xml
 - The application is installed on the Datacap Server.
 - Upload batched with the Upload Batch Profile
- At a remote location only accessible over the internet.
 - Application Manager > Service * = C:\Datacap\datacap.xml
 - The copy of the Server application is installed on the remote FastDoc workstation.
 - Upload batched with the Web Client Upload Service.



- The Application Manager Service * parameter sets the application location.

Figure 2-21. Two FastDoc workstation configurations

On the same LAN as the Datacap Server.

- The Application Manager > Service Tab > “Path to the application manager file” parameter, is set to \\<Server>\Datacap\datacap.xml
- The Application and Workflow parameters reference an application that is configured on the Datacap Server.
- Upload batches to the Datacap Server with the Upload Batch Profile.
- The offline configuration Task Option “Send to Datacap” has a Batch Folder parameter.
 - If the Batch Folder parameter is set, then upload batches go to the location on the server that the Batch Folder parameter specified.
 - If the Batch Folder parameter is not set, then the upload destination is determined by concatenating the following entities. Application Manager > Service “Path to application manager file” + Application that is selected on the Send to Datacap task option + batches.

Example: \\<Server>\Datacap\<application>\batches

At a remote location only accessible over the internet.

- The Application Manager > Service Tab > “Path to the application manager file” parameter, is set to C:\Datacap\datacap.xml
- The Application and Workflow parameters reference a copy of the server application that is installed in the remote FastDoc workstation.
- Upload batches to the Datacap Server with the Web Client Upload Service.

Create and run local batches

- Local Mode

- Click a batch profile name under New Batch to create a batch.
- Double-click an Existing batch.



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-22. Create and run local batches

Local Mode

- In the local mode, FastDoc runs all defined tasks automatically until a step is reached that requires operator intervention.
- The results of the FastDoc Local processing are presented in one of two ways dependent on the configuration of the Local tasks.
 - Complete processing is done under FastDoc and the results are placed on the destination repository as defined in the FastDoc Export step.
 - FastDoc batches are on the client system where FastDoc is installed at C:\Datacap\FastDoc\Batches\<Batch Profile Name>\<Batch Folder>. At the point that the Send to Datacap Task Option is encountered, a Batches Folder is created on the Datacap server system where the Datacap Batches are stored at <Datacap Server>\Datacap\<Application>\Batches\<Batch Folder>
- At any time, an operator can place the batch on hold mode. An entry is made on the Existing Batches list so that the batch can be restarted later.

Exercise: FastDoc local mode

Requires:
Course Exercise Guide
Student system

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-23. Exercise: FastDoc local mode

Exercise objectives

- Explore FastDoc local mode.

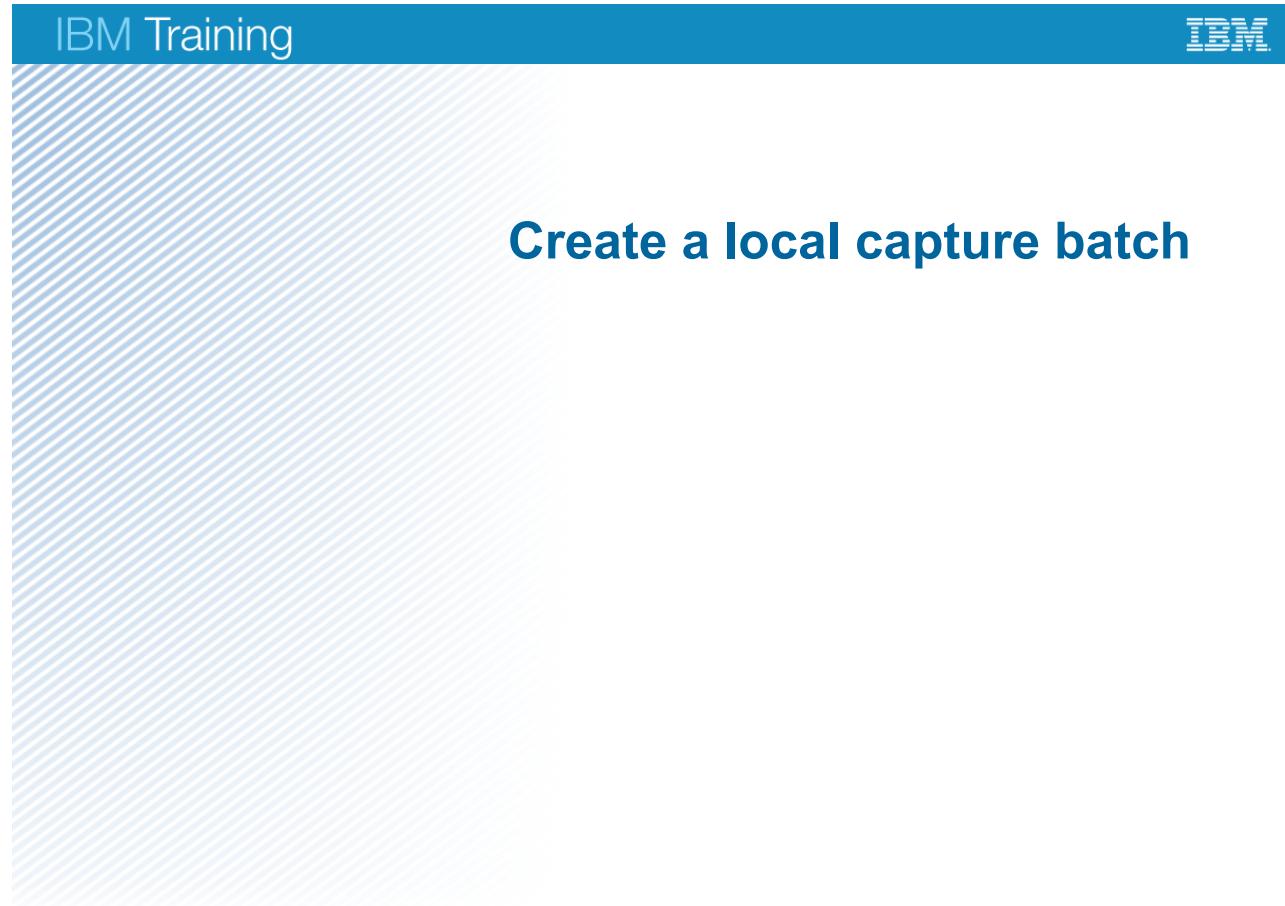


FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-24. Exercise objectives

Lesson 2.3. Create a local capture batch



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-25. Create a local capture batch

Lessons

- FastDoc Overview
- FastDoc local mode
- Create a local capture batch
 - Process local batches
 - Configure Scan and Upload

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-26. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. You might be called on to configure remote site for scanning and uploading document batches to the Datacap Server.
- To do this operation, you must be familiar with using the FastDoc user interface in the local run mode to capture documents of a known type.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-27. Why is this lesson important to you?

Define terms

- Local refers to the machine on which FastDoc is installed.
- FastDoc Client Workstation refers to the workstation where FastDoc is installed.
 - It might be on the same local area network as the Datacap Server.
 - It would normally be run in Datacap Server mode.
 - It might be at a site remote from the Datacap server.
 - It would normally be run in local mode.
- Local or Offline mode
 - Refer to a mode where capture tasks are run on the machine where FastDoc is installed.
 - At some point, the local batches are uploaded to the Datacap Server.
 - FastDoc can be used to scan directly to a repository without ever uploading to a Datacap Server.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-28. Define terms

FastDoc might be used with the Datacap Client Upload Service to upload to a Datacap Server at scheduled times or when the server is available. Scheduled upload times are useful in WAN situations with low or expensive daytime network availability.

What is a Local Batch Profile?

- It is a collection of tasks that are used to process a batch on the local machine.
- The batch profile on the configure page defines:
 - How to name the batch.
 - How to identify sequential batches.
 - How to name captured page files.
- A local batch has a defined task for each step that is run on the local machine.
- If all processing is on the local machine, the last task exports the documents to a repository.
- If a batch is only partially processed in the local mode, then the last task uploads the batch to the server for further processing.

Figure 2-29. What is a Local Batch Profile?

- The batch data is maintained on the local FastDoc machine until the batch is uploaded to the Datacap Server.
- The batch data location on the local machine is: C:\Datacap\FastDoc\<appname>\batches

What is a Local Task?

- The tasks that are defined in the batch profile point to profiles already configured in an existing application.
 - There is a task defined for each task that is run on the local machine.
- Each task is either defined as:
 - A Verify task by selecting the Verify task option.
 - Example: Verify task
 - A Profiler task by selecting the Profiler task option.
 - Examples: PageID, Profiler, Export.
- The application might be on:
 - The Datacap Server.
 - The Local Client Workstation

Figure 2-30. What is a Local Task?

- The task that does the upload operation defines where the batch is placed on the Datacap server.
- The Send to Taskmaster task option defines the server batch destination. This topic is covered in the next lesson.
- If no path is defined in the Batch Folder field of the Send to Taskmaster option, then the path that is defined in the Application manager is used.
- At the point of uploading the batch to the Datacap server, the local copy becomes redundant.
- It is an administrative task to monitor and remove the redundant batch data before it uses up the FastDoc client machine storage resources.

Location of Application that is run in FastDoc Local mode:

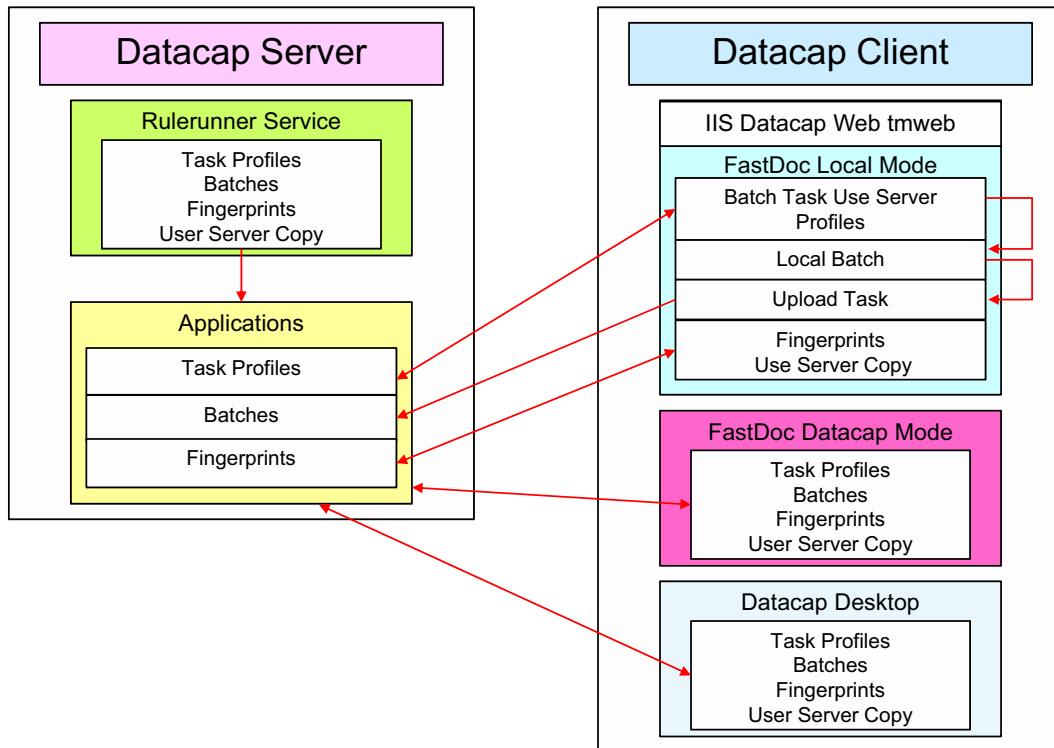
- If the Client is on the same LAN as the server, then the application that the Local Batch Profiles references is on the Datacap Server.
- If the Client is on remote LAN from server with only WAN access to the server, then the application is on the client or another local machine.

When is a Local Batch Profile required?

- A local batch profile is required when any capture processing is planned on the FastDoc client machine in the Local mode.
 - When a local batch profile is built for an application; it should not be necessary to access it again.
 - The only possible causes for changing the initial configuration for a specific application are:
 - If you want to adjust the file and folder name configuration parameters.
 - If processing for the application changes so that another task is added.

Figure 2-31. When is a Local Batch Profile required?

Where is data located?



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-32. Where is data located?

The slide diagram shows where task profiles, batches, and fingerprints are stored.

Server Object Location

Objects - Where located

Datacap Server - ecmedu01

Applications - \\ecmedu01\Datacap\<application>

Batches - \\ecmedu01\Datacap\<application>\batches\<batch folders>

By default when Datacap Desktop, Rulerunner, Datacap navigator, and FastDoc when it is run in Datacap Server mode, create batches they are placed here.

Also uploaded batches from tmweb and FastDoc offline mode can also be placed here.

Fingerprints - \\ecmedu01\Datacap\<application>\fingerprints

Task Profiles - \\ecmedu01\Datacap\<application>\dco_<application>

Datacap Client - dcclient

FastDoc - \\dcclient\Datacap\FastDoc

Local Batches - \\dcclient\Datacap\FastDoc\batches\<application>\<batch folders>

Fingerprints - Use the Server Fingerprints

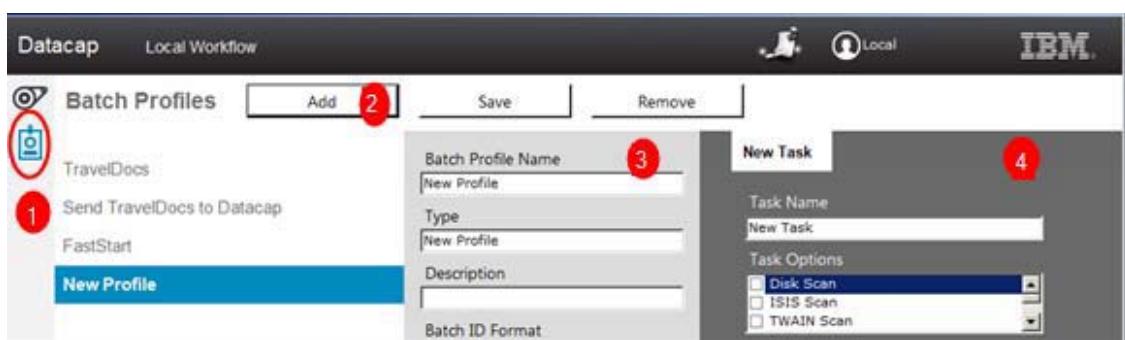
Local Batch Profiles - \\dcclient\Datacap\FastDoc\BatchProfile.xml

All configuration that is done on the setting window is stored in this file.

These Local Batch Task objects reference the task profiles on the server.

IBM Training

Configure local batches



- Select the Configure local profile view.
- Use Add to create another batch.
- The Batch Profile Name defines the Batch tab name.
 - The Batch Profile Name is also used to create a batch folder structure:
 - C:\Datacap\FastDoc\<Batch Profile Name>\batches\<date>.<num>
- Use Remove to delete a batch.
- Use Save to save any changes.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-33. Configure local batches

Configure local profiles view toolbar

- Add – Add an offline batch
- Save – Save configuration changes
- Remove – Remove a defined local profile

Left pane

- List of batch profile links

Center light gray pane

- Batch profile parameters

Right dark gray pane

- Task configuration options and parameters

Batches and Tasks

- Applications can be configured to recognize and process multiple different documents.
- Usually you create a batch profile for each application processed.

- There might possibly be multiple batch profiles that are defined for an application.
 - Example: One to process all tasks offline. One to Scan and then Upload.
- A batch profile defines all of the tasks that must be done before the batch is uploaded to the Datacap server
- You can do any number of tasks in the Local mode before sending the result to Datacap server.
- As soon as the Send to Taskmaster Task Option is encountered in a task, it signifies the end of the local processing.
- All remaining tasks are done on the Datacap server.
- FastDoc and any other Datacap client can be used to process the tasks on the Datacap server.
- The Rulerunner server can be configured to run any task that does not require intervention.



FastDoc (Admin) local configuration interface



- There is normally only one batch profile per application.
- One batch link for each defined batch.
 - One click the batch link causes all of the associated tasks to run.
 - The batch profile name defines the batch Shortcut name.
- Order of processing
 - They run in the order defined. Scan, Profiler, Verify, and Export.
 - Processing pauses only when user action is required. Scan and Verify

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-34. FastDoc (Admin) local configuration interface

One batch profile per application

- There is normally only one batch profile shortcut for each application that is configured to process in the local mode.
- The reason for a single shortcut per application is that all tasks run automatically after the batch scan task is initiated.

Local batch profiles are configured on the Configure local profile view.

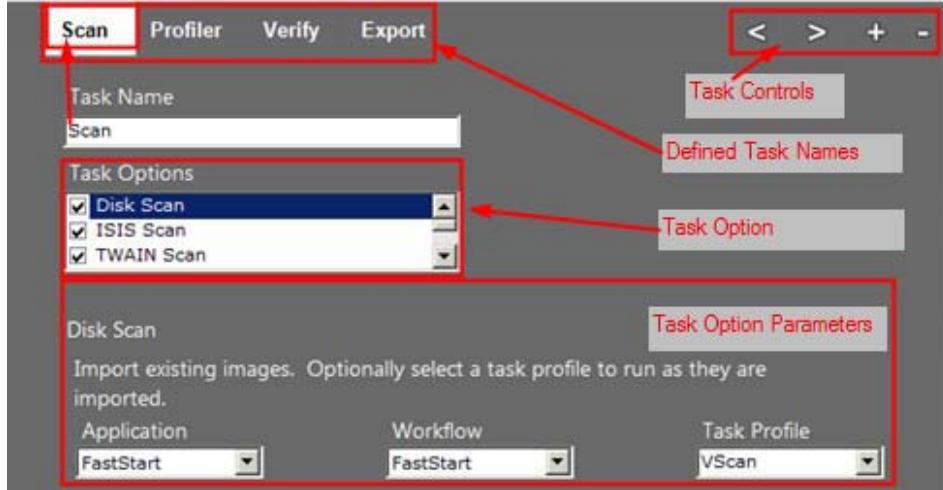
- Configure local profile view can be accessed only from the FastDoc (Admin) interface.
- The administrator configures the local profiles for the users.
- Configuration that is done on the Configure local profile view defines local batch profiles.
- Users run the FastDoc (User) interface in the local mode.
- No Datacap enforced security prevents a user from running FastDoc local mode.
- The Administrator can set security on the C:\Datacap\FastDoc folder to limit access.
- The FastDoc batch profiles can be duplicated across FastDoc client machines by copying the file C:\Datacap\FastDoc\BatchProfiles.xml.

Difference between processing Online and Offline

- The location of the batch temporary files.
 - Local Batches - \\<local machine>\Datacap\FastDoc\batches\<application>\<batch folders>
 - Batches - \\<datacap server>\Datacap\<application>\batches\<batch folders>



Batch task pane



- Task Name is displayed on the toolbar.
 - There is a tab for each task that you define.
 - Only the batch name is visible to user at run time.
- Task toolbar controls, add, remove, or sequence batch tasks.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-35. Batch task pane

The order in which tasks are arranged determines the order of processing.

Sequencing batch tasks.

–< > Left and Right. Use these controls to order the task menu tabs.

–+ Plus sign to add a task.

-- Minus sign to remove a task.

Task Options

These options are the configuration options available for configuring the tasks.

- Disk Scan (optional, required if scanning digital images)
 - Import existing images.
 - Run the selected task on all images as they are imported. This task might be used for converting documents to TIFF.
- ISIS Scan - Scan batches from an ISIS scanner device.
- TWAIN Scan - Scan batches from Twain scanned device.

- Start Batch Panel
 - This option is used with the Scan options to cause the Start Batch Panel to be displayed to control importing.
- Verify - Identify a Verify task.
- Rulerunner
 - Identify Rulerunner tasks. (PageID, Profiler, and Export)
- Send to Taskmaster
 - The batch folder setting under upload is optional and if omitted the one in the application file is used.

Configure more parameters

- Each task option has a unique set of required parameters.

Task Options and associated Parameters

In this topic, the first-level bullet shows the options that can be selected for each task type and the second-level bullets (-) define the parameters that can be selected for each option.

Scan Task Configuration Options

- Disk Scan (optional, required if scanning digital images)
 - Application = <app name>, Workflow = <workflow name>, task = VScan (optional)
- ISIS Scan (optional, required if scanning with an ISIS scanner)
- TWAIN Scan (optional, required if scanning with a TWAIN scanner)
- Start Batch Panel
 - Application = <app name>, Workflow = <workflow name>

Profiler Task Configuration Options

- Rulerunner
 - Application = <app name>, Workflow = <workflow name>, task = Profiler

Verify Task Configuration Options

- Verify
 - Application = <app name>, Workflow = <workflow name>, task = Verify

Export Task Configuration Options

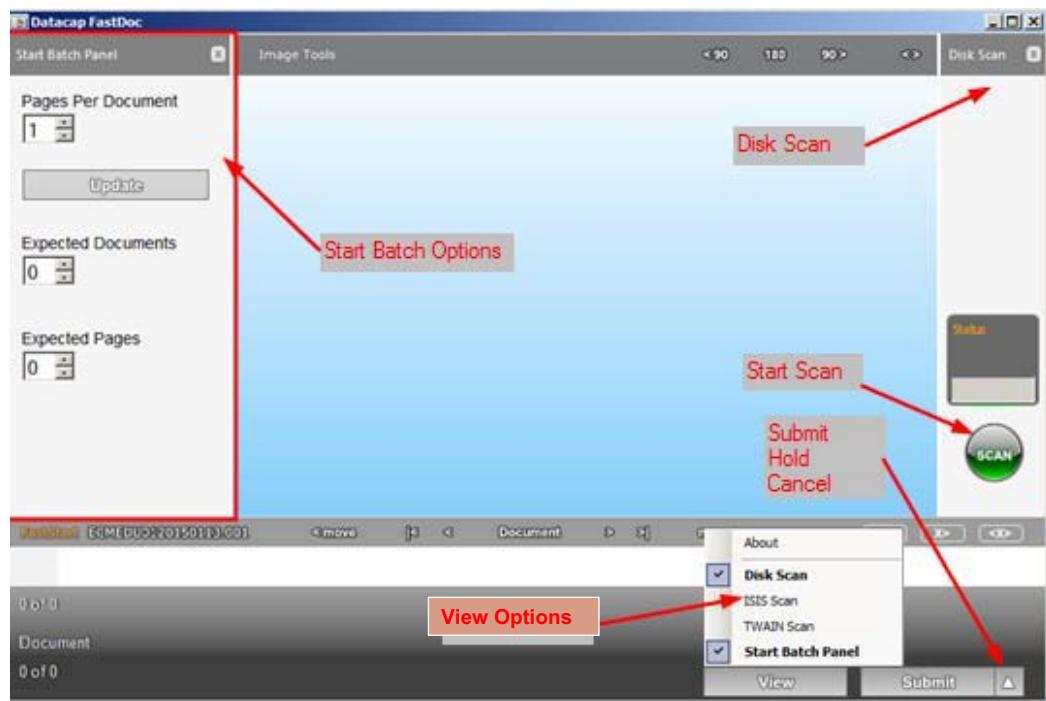
- Rulerunner
 - Application = <app name>, Workflow = <workflow name>, task = Export

Upload Task Configuration Options

- Send to Taskmaster
 - Application = <app name>, Job = Main Job, Statistic Table = scan2kst
 - Batch Folder = \\ecmedu01\\Datacap\\ExpenseDemo\\batches

IBM Training

Scan Interface



- You see the Scan interface when you start scanning.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-36. Scan Interface

Introduction

After the Batch profile and Task configuration is complete, you can start processing document batches with FastDoc. You see a similar interface when processing batches in Local - Offline mode or Taskmaster - Online mode.

Start Batch Panel option

The following options help to separate batches of pages into documents.

- Pages per Document

If all documents have the same number of pages, then quickly use this option to group pages into documents.

- Update

Separate pages into documents with the Pages per Document value.

- Expected Documents

The following option specifies the total number of documents in the scan batch.

- Expected Pages
 - This option specifies the total number of pages in the scan batch.
 - Note: The Expected documents and pages values are saved to batch variables. Rules can be used to route the batch for fixup if the actual numbers do not match.

View options

- Used to access Help, Access information about the software release, and the select the type of Scan panel that is displayed of the right side pane.
 - Help - Display Help information.
 - About - Display information about the software release,
 - Disk Scan - Default Scan interface that is displayed on the right side pane.
 - ISIS Scan - Select this option to display the ISIS Scan interface on the right side pane.
 - TWAIN Scan - Select this option to display the TWAIN Scan Interface on the right side pane.

Note: The Disk Scan, ISIS Scan, and TWAIN Scan are only displayed on the View list if the check box for these options were selected when configuring the Batch Scan Task. See the Scan Task Configuration Options on the previous slide.

Completion actions

- There are three possible completion actions:
 - Submit - Complete the Scan task and allow processing to continue to the Profiler task.
 - Hold - Suspend the batch at the Scan step.
 - Cancel - Cancel the Scan and all subsequent tasks.



Scan Panels



- The scan interface looks different depending on scan source option that is selected.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-37. Scan Panels

Scan interfaces for Disk Scan, ISIS Scan, and TWAIN Scan.

- Disk Scan - Scan Creates a batch from documents that are previously scanned or faxed.
- ISIS Scan - Select Source Creates a batch from images that are scanned with an ISIS scanner.
- TWAIN Scan - Select the TWAIN source Creates a batch from images that are scanned with a TWAIN scanner.

ISIS TWAIN Differences

- ISIS scanner drivers provide consistent functions across a wide variety of scan devices.
- TWAIN scanner drivers provide more specialized functions.

Exercise: Create a local capture batch

Requires:
Course Exercise Guide
Student system

Figure 2-38. Exercise: Create a local capture batch

Exercise objectives

- Configure a local batch.



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-39. Exercise objectives

Lesson 2.4. Process local batches

Process local batches

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-40. Process local batches

Lessons

- FastDoc Overview
- FastDoc local mode
- Create a local capture batch
-  Process local batches
- Configure Scan and Upload

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-41. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. You might be called on to configure remote site for scanning and uploading document batches to the Datacap Server.
- To do this operation, you must be familiar with using the FastDoc user interface in the local run mode to capture documents of a known type.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-42. Why is this lesson important to you?

Scan batch preparation

- Scan task is the first step in Document capture
- Batch content
 - Physical paper is scanned with an ISIS or TWAIN scanner
- Electronic documents that are received through:
 - Fax, email, previously scanned images.
- Scan preparation
 - Documents match the Datacap application
 - For paper batches:
 - Arrange page groups for multi-page documents.
 - Add separator pages if necessary
 - For electronic batches, know the location:
 - Single folder
 - Multiple batches from different folders

Figure 2-43. Scan batch preparation

- The Scan Task is the first step when capturing and processing batches of documents.
- Batch content might be physical paper pages that are scanned with a scanner.
- Content might also be pages that are received through:
 - Fax.
 - Email.
 - Images that were previously scanned and placed in a specific image folder for Datacap processing.
- Make sure that documents are the correct type for the application that is used to scan them.
- If scanning paper documents:
 - Arrange pages into document groups if multi-page documents.
 - Add document separator pages if necessary.
- If processing graphic images, make sure that you know the location.
 - In one folder
 - Multiple batches from different folders



Start a batch in local mode

- Start either FastDoc (Admin) or FastDoc
 - Click Local and Login.



- Click the batch name
 - (for example, ExpenseDemo) for the batch to run.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-44. Start a batch in local mode

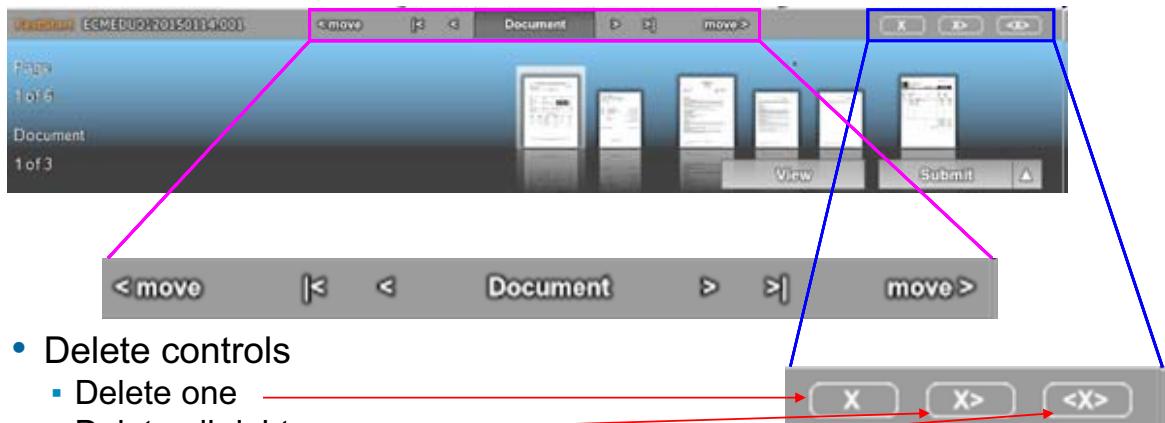
Fast Start Application

- For the exercise, you use the ExpenseDemo application.
 - It was created from standard DC8.1 FastApp application template
- The application is configured with four basic tasks.
 - Scan, Profiler, Verify, and Export
- It is configured to recognize three document types
 - Fingerprints were saved for each document type.
- Location rules were defined for locating and extracting data.
- Some validation rules were configured to do field data checks.
- Local batch tasks are configured to use the same task profiles as online processing.



Scan the batch

- Scan to start scanning
- Document controls and image strip



- Delete controls
 - Delete one
 - Delete all right
 - Delete all

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-45. Scan the batch

After the Scan interface opens:

- Click Scan to select the images for scanning.
- Browse to and Select one or more images.
- Click Open to start the scan process.

Thumbnail images of each scanned page are displayed on the thumbnail strip.

- The first image is placed in center of the blue background pane of the Scan window.
- The first image is highlighted on thumbnail strip.

Move and select image controls

Use **<move and move>** to move the selected page left or right.

Use **|< and >|** Go to first page and Go to Last page.

Use **< and >** Go to Precious and Go to Next

Use **Document** to define the selected page as the first page of a document. The tile size increases.

Delete controls

Delete only the selected image

Delete the selected image and all images to the right of the selected.

Delete all images.

Process the batch

- Start profiler.
 - Click Submit on the Scan window to initiate the Profiler task.
- Processes that the profiler task performs are:
 - Fixing images
 - Recognition
 - Locate the Data
 - Populate the Data
- Fields that fail validation rules are flagged for operator review during the verify task.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-46. Process the batch

Completing the scan task initiates the profiler processing task.

- Click Submit to complete the Scan task after doing all necessary document manipulation.
 - In Local mode, the Profiler task or next sequential task starts to run.
 - In Datacap Server mode, Scan task ends and Profiler or next sequential task must be run manually.
 - Problems are identified in the profiler task and flagged for the operator to review when the verify task is run

All processing that is done in profiler task is part of standard processing that is built in to the application template.

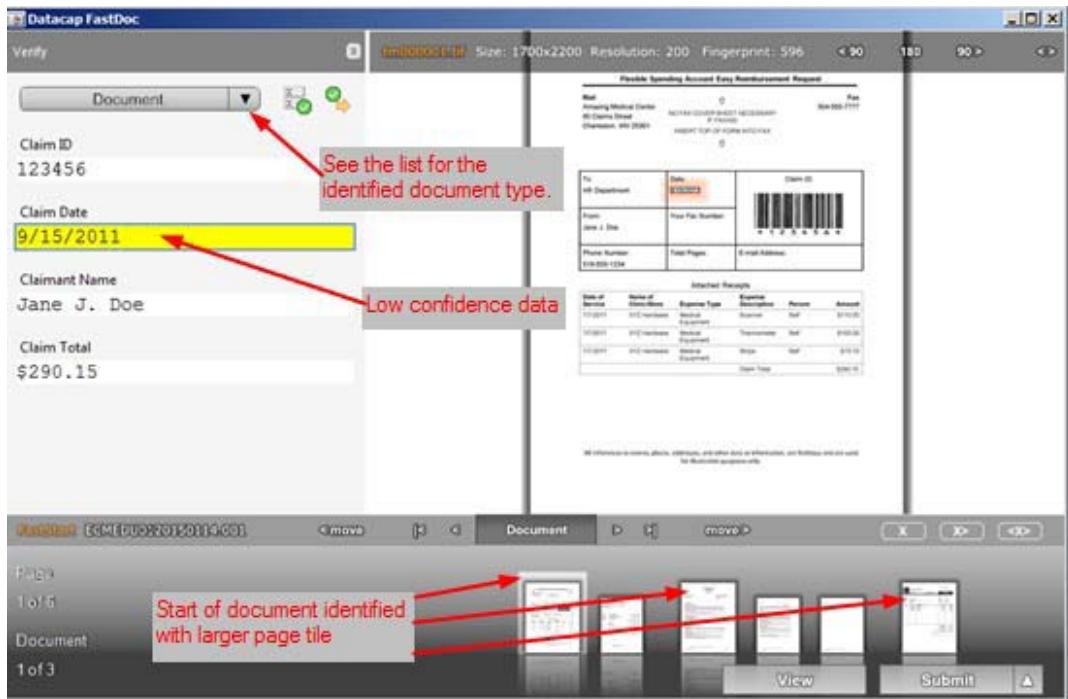
Some additional basic configuration is done for the three document types that are processed in this lesson.

The following are the processing phases for the Profiler task:

- Fixing images
 - Do image alignment, despeckling, and so on, to improve image quality.
- Recognition
 - Identify pages by comparing them to saved fingerprints.
- Locate the Data
 - Use configured locate rules to locate and extract image data fields.
- Populate the Data
 - Populate verify form fields with extracted data.



Verify the batch



- When the Profiler completes, the Verify window is displayed.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-47. Verify the batch

The verify interface is where Quality assurance of the scanned data occurs.

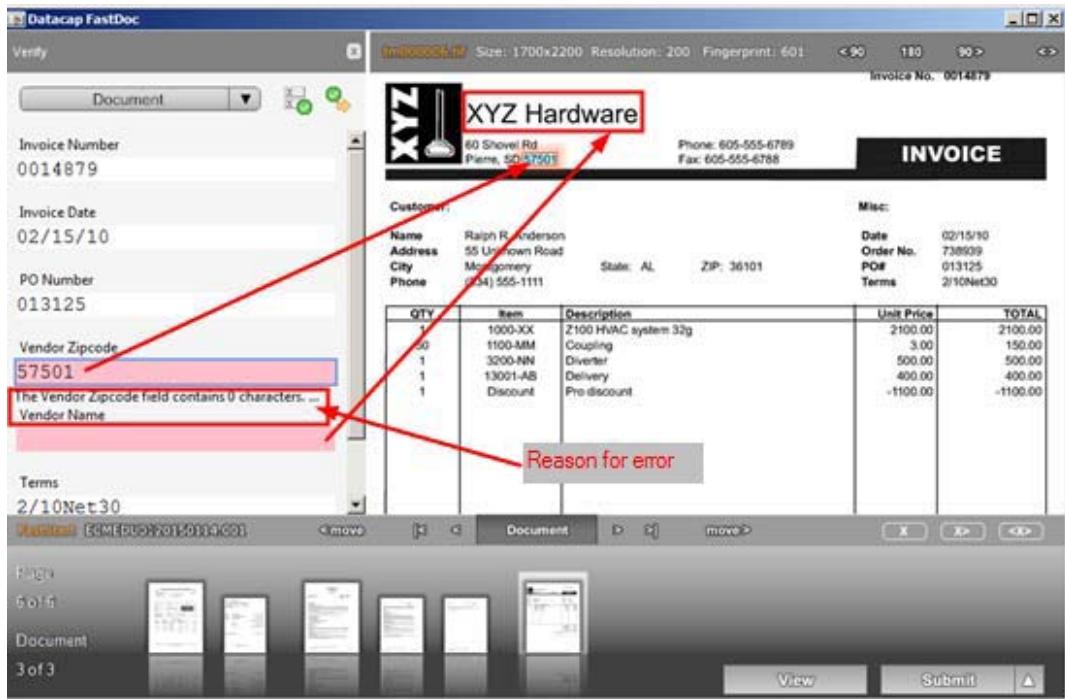
When profiler processing is complete, the Verify window is displayed.

- The Verify window opens with:
 - A tabular Verify panel that displays the extracted values.
 - The full image in the right side pane.
 - The Thumbnail strip across the bottom.
- Verify that the Window Data Panel displays the current document.
 - It was identified with a previously defined fingerprint.
- Extracted properties are displayed in verify/data entry screen in a tabular layout.
 - The Claim Date field has a yellow background that indicates low confidence data.
 - Fields without data are displayed with a pink background.

- The Thumbnail Strip is across the bottom of the Verify window.
 - Pages are grouped into documents. The first page has a larger icon.
 - First document that has a flagged field is highlighted in the page thumbnail strip.
 - You can select a page of the batch by clicking its thumbnail.



Correct errors



- Complete missing fields with click'n'key.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-48. Correct errors

- The Verify task flags potentially problematic fields based on
 - character interpretation,
 - confidence,
 - data validation rules,
 - empty data fields, and so on
- Missing data fields display with a pink background.
 - The message after the field displays a reason for the verify failure.

Provide Missing Data or Correct Erroneous Data

- Click'n'Key Data retrieval.
 - Select the empty field in the Verify Data panel.
 - For a single word or contiguous data field, click the corresponding field on the image.
 - If the field on the image has multiple words, then draw a box around the area to select all of the words.

- Alternatively, type in missing data for fields that need to be corrected.
- Note: Use keyboard shortcuts to zoom or pan
- After the user provides or corrects data, the changed data is revalidated.
- If user submits a page, the interface advances to the next page with a problem.

Export the batch

- Validate Changed Data
 -  Run the validation rules after providing missing data or correcting fields.
- Submit Pages
 -  Submit changes advances the process to the next page that has a suspected problem.
- The export task runs automatically after the last verify problem is resolved
- At the point when export completes on local machine,
 - Batch is in C:\Datacap\FastDoc\ batches\<app name>\<batch folder>
- After Upload, batch is transferred to server and placed in
 - \Network\<Datacap Server>\Datacap\<app name>\<batch folder>

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-49. Export the batch

When you click Submit after the last problem is resolved:

- In the Local mode, the verify task ends and the Export task starts.
- In the Datacap Server mode, the Verify step completes and the batch is set as pending on the Export task.

The Export files are in the batches folder because no Export configuration is done on the ExpenseDemo application.

In a later exercise you configure the Export files to be placed in:

- A specific Exports folder on the Datacap Server machine.
- The Enterprise Content Manager (ECM) repository on the IBM FileNet P8 Content Engine.

Exercise: Process local batches

Requires:
Course Exercise Guide
Student system

Figure 2-50. Exercise: Process local batches

Exercise objectives

- Process documents in local mode.

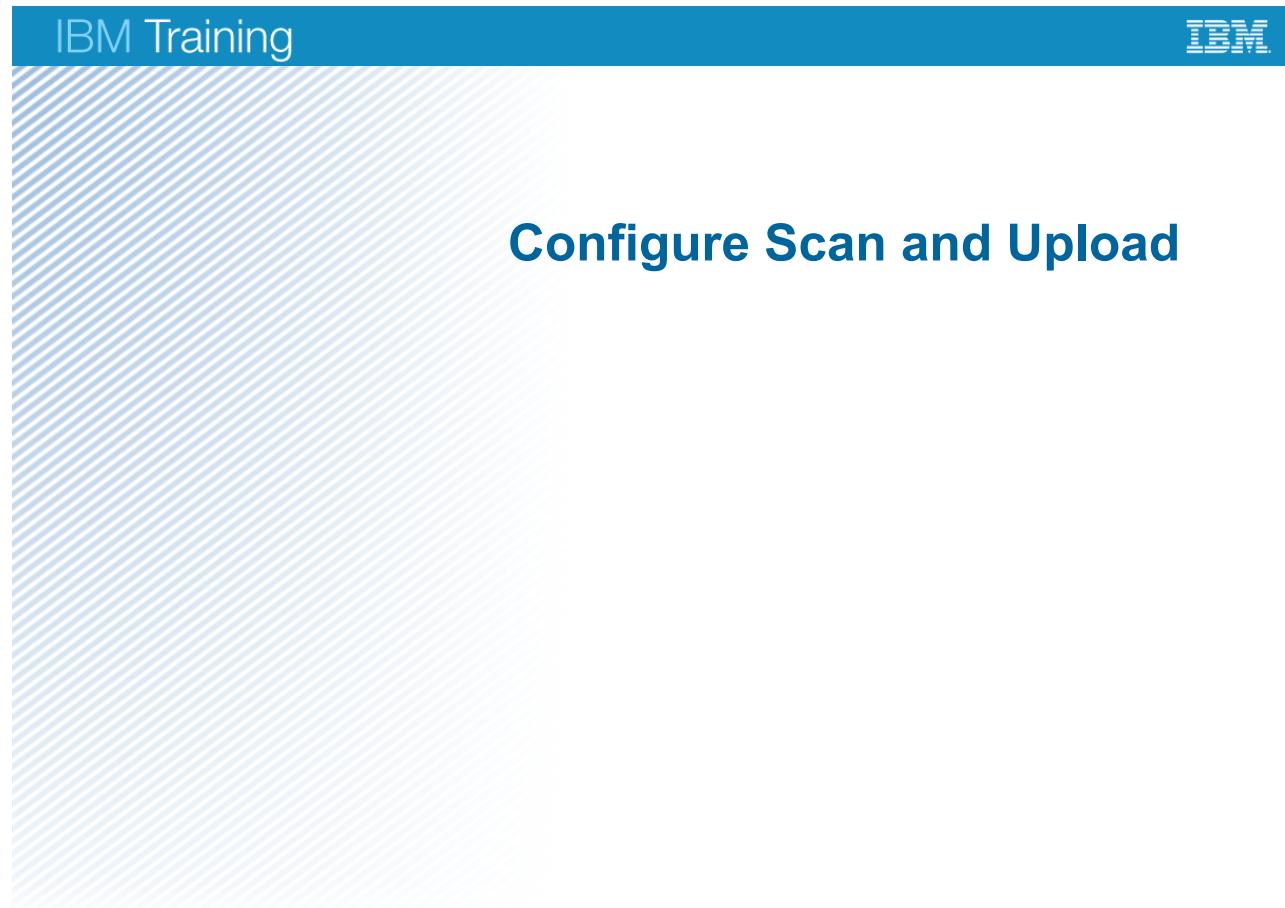


FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-51. Exercise objectives

Lesson 2.5. Configure Scan and Upload



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-52. Configure Scan and Upload

Lessons

- FastDoc Overview
 - FastDoc local mode
 - Create a local capture batch
 - Process local batches
-  Configure Scan and Upload

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-53. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. You might be called on to configure remote site for scanning and uploading document batches to the Datacap Server.
- To do this operation, you must be familiar with using the FastDoc user interface in the local run mode to capture documents of a known type. You must also know how to configure an Upload shortcut with FastDoc.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-54. Why is this lesson important to you?

Add an Upload Batch

- When is an upload task required?
 - When some processing is local and the rest is central.
 - If you want to configure FastDoc as a remote scan station.
- The upload task is created in the FastDoc (Admin) Local mode.
 - Start the FastDoc (Admin) interface.
 - Select local mode and log in.
 - Click Add to define a new batch.
 - Configure the Batch profile parameters. (Covered in lesson 1.3)

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-55. Add an Upload Batch

Add an Upload Task

When to use

- An Upload Task is required in any Scenario where you do some processing in the FastDoc Local mode and then upload the batch to the Datacap server to complete the processing.
- If you did all of your processing on the local workstation, then you would typically upload following the Export task.
- By convention upload happens at the end of the last defined Local batch task.
- The remainder of the processing is done on the Datacap server.

Before you begin

- You must start the FastDoc (Admin) interface to define new batch tasks.
- You can start FastDoc (Admin) in the Local - Offline mode.
- Because you are configuring the upload task in the Local – Offline mode, you are forced to provide login credentials when you define the Send to Datacap Server step.

Batch profile tasks configuration options

- Where you configure local batch tasks
- On this pane you configure:
 - A task for each task that must run in the FastDoc machine.
- For each task you configure:
 - Task name
 - Task options
 - Application to connect to
 - Workflow name
- Arrange tasks in order of processing.
- FastDoc runs tasks in the order that they are defined.

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-56. Batch profile tasks configuration options

What is a batch profile task?

The batch profile task interface is where you configure batches and tasks that are used for local batch processing with FastDoc. The batch and task configuration is done on the Settings window.

- A term batch in this context is group of tasks that define the processing that is done when documents are captured in the offline mode.
- Not to be confused with the term batch that refers to a group of documents that are all processed at the same time.
- The Batch profile task defines the task name, options, which application, and workflow are used to process the task. Examples: Scan, Profiler, Verify, and Export.
- A Batch profile task is defined for each task that is run on the batch in the Local - Offline machine.
- Batch profile tasks must be arranged in the order in which they are run.
- When the batch is initiated, FastDoc processes all of the tasks in the sequence in which they were defined.

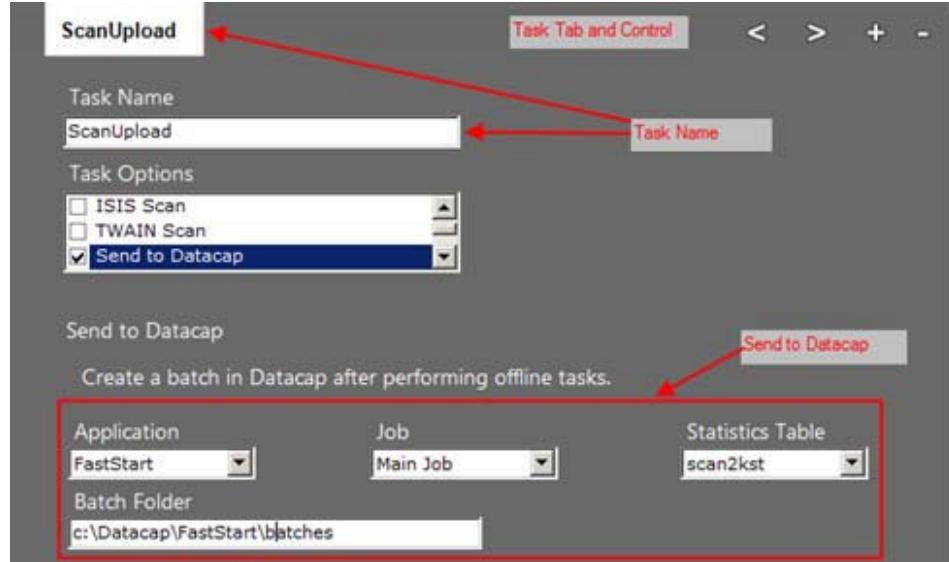
Batch Tasks Profile Examples:

These tasks names are examples of task that might be included in a batch.

- Scan - defines the scan task for import of electronic images from a physical scanning device.
- Profiler - defines image cleanup, page recognition (OCR), and validation tasks.
- Verify - provides an interface that is used to review and update field data.
- Export - moves documents to a supported repository.
- Upload - Transfers batches to the Datacap server.



Configure the Upload Task



- Name the task. Example (ScanUpload)
 - Configure DiskScan option
 - Configure Send to Disk option
 - Configure Start Batch Panel option

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-57. Configure the Upload Task

In the Task pane type, enter the Task Name. Example: ScanUpload.

Configure the Scan Task options.

- Select the Disk Scan option in the Task Options pane.
- Select the Application and Workflow. Example ExpenseDemo

Configure the Send to Taskmaster Task options.

- Select the Send to Taskmaster Task option in the Task Options pane.
- When you select the Send to Taskmaster option, you are prompted to provide log in credentials.
 - Select Application: ExpenseDemo, Job: Main Job, and Statistic Table: scan2kst.
 - The Batch Folder is the folder on the Datacap server where the batch should be placed. Example \\ecmedu01\\Datacap\\ExpenseDemo\\batches
 - If the batch folder is left blank, the folder that is specified in the application manager is used.

Configure the Start Batch Panel Task options.

- Select the Start Batch Panel option in the Task Options pane.
- Select Application and Workflow. Example ExpenseDemo. Optionally select a Task Profile Scan.

Exercise: Configure scan and upload

Requires:
Course Exercise Guide
Student system

FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-58. Exercise: Configure scan and upload

Exercise objectives

- Configure a scan and upload task.



FastDoc Local Mode

© Copyright IBM Corporation 2016

Figure 2-59. Exercise objectives

Unit summary

After completing this unit, you should be able to:

- Identify the operating modes of Datacap FastDoc.
- Configure batches for off line processing.
- Scan and upload a batch

Unit 3. FastDoc Datacap Server mode

Overview

This course provides an overview for using Datacap on the Server mode. You create an application from the FastDoc application Form template using the Forms template.

Objectives

- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacapoc.doc/datacap_9.0.1.htm

Unit Objectives

- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-1. Unit Objectives

Lesson 3.1. Introduction to FastDoc Datacap Server mode

Introduction to FastDoc Datacap Server mode

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-2. Introduction to FastDoc Datacap Server mode

Lessons

Introduction to FastDoc Datacap Server mode

- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-3. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you must often quickly scan new data batches without having to go through a long development cycle.
- To do these tasks effectively, you must be familiar with the FastDoc Datacap Server mode interface.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-4. Why is this lesson important to you?

FastDoc Datacap Server interface

- Run FastDoc as a Datacap Client. (Datacap Server mode)
 - Scan, index, and run background tasks on documents and image files.
 - This capability is in either the User or Admin Interface.
- Run FastDoc as a rapid application development tool.
 - Use Application Wizard to create Applications from default templates.
 - Standard Jobs and Tasks are defined for you in the templates.
 - Define the Batch Structure (DCO)
 - Configure workflows for capture, recognition validation, and export.
 - Define fingerprints.
 - Use the Test feature to verify that location and recognition are working as configured.
 - This capability is only available in the Admin interface.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-5. *FastDoc Datacap Server interface*

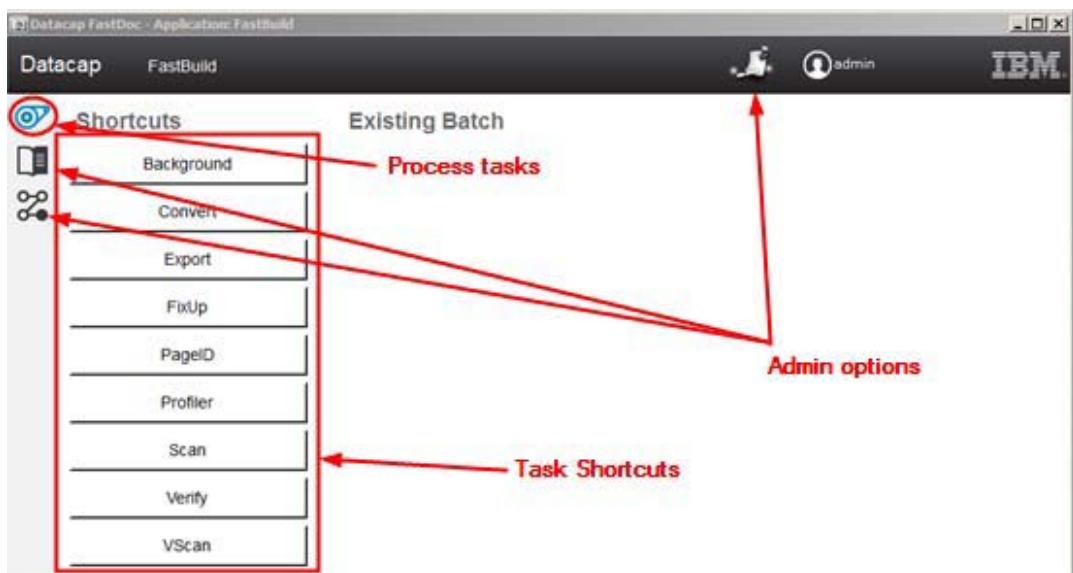
FastDoc is a client that you can use to scan, index, and manually run background tasks on documents and image files.

- You can run FastDoc in a stand-alone environment, where FastDoc manages its own batches. Stand-alone mode was covered in the previous unit.
- You can run it as a client to the IBM Datacap Server in place of Datacap Desktop.
- You can also use FastDoc as a rapid application development tool.

IBM Training



Introduction to FastDoc Datacap Server mode



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

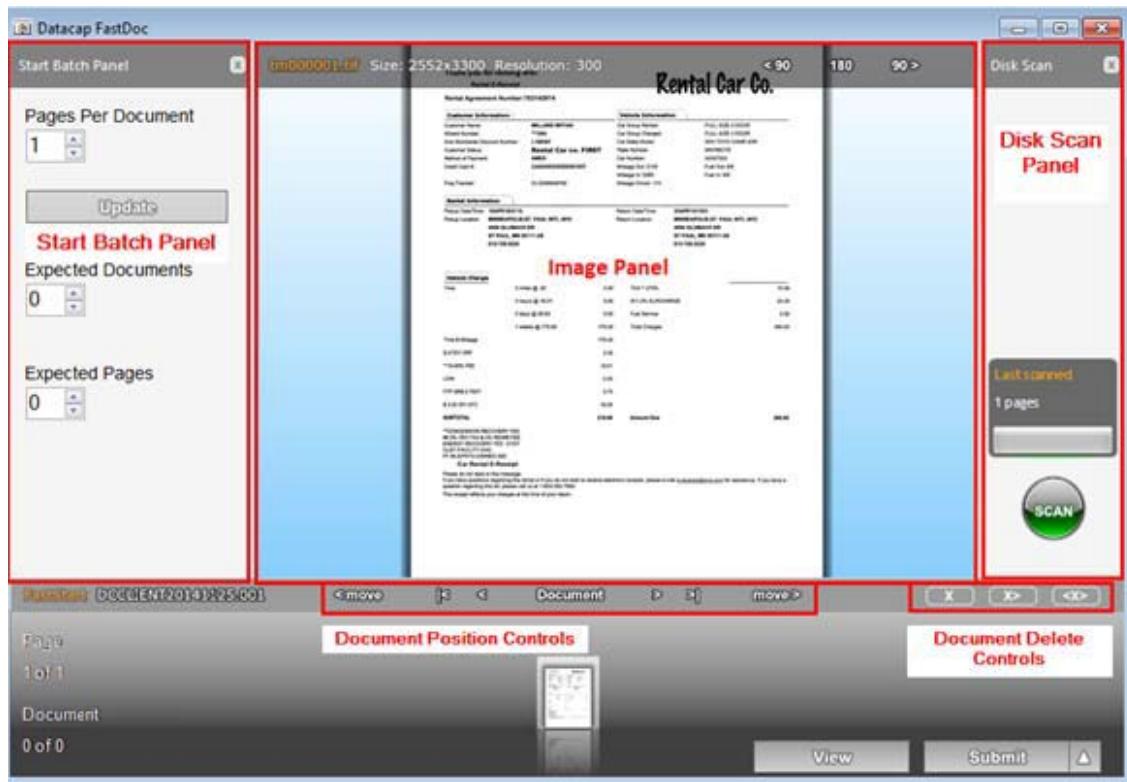
Figure 3-6. Introduction to FastDoc Datacap Server mode

The slide image represents the FastDoc online client time interface.

- Datacap Server mode is the online mode.
- The application name is in the heading of the interface.
- Shortcuts reference task profiles in the application.
- Unsupervised tasks – Profiler and Export – don't run automatically as in local mode.
- Therefore, it is necessary to have a task profile link for each so they can run manually.
- Alternatively, tasks that do not require operator intervention like PageID, Profiler, and Export, can be run unsupervised if they are configured to run in Rulerunner.



Scan Task Interface



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-7. Scan Task Interface

The Scan Task Options

- Disk Scan (For processing digital images. Email, Fax, or digital images that were previously captured and stored on the disk.)
- ISIS Scan (For ISIS compatible scanners).
- TWAIN Scan (For TWAIN compatible scanners).
- Start Batch Panel. (To display the Start Batch Panel when scanning).

Scan task Interface

- Left Pane – Start Batch Pane. You enabled the Start Batch Panel with:
 - Start Batch Panel Task Option for the Offline mode.
 - tmweb > Administrator > Workflow > <Jobx> > iVScan, VScan, or Scan tasks > Setup FastDoc pane > Show the Start Batch Panel
 - Page Structure controls.
 - Pages Per Document (To automatically separate documents when all of the documents have the same number of pages.)

- Expected Documents (The operator sets the expected number of documents in the batch and it checked programmatically during processing.)
- Expected Pages (The operator sets Expected number of pages (images) to limit the number of pages that are read at scan time.)
- Right pane – Scan type Panel. You select what is displayed in the Scan Type Panel from the View menu. The Task Option selections define what populates The View menu.
 - Disk Scan (For processing previously scanned or imported digital images.)
 - ISIS Scan (For scanning paper batches by using an ISIS compatible scanner.)
 - TWAIN Scan (For scanning paper batches by using a TWAIN compatible scanner.)
- Bottom pane with the gray background, Thumbnail images for each scanned page.
- Center pane – Selecting a thumbnail image causes it to be displayed in the Image pane.
- Thumbnail Control Bar.
 - Document Position Controls, you can select the thumbnail and position it to get pages in the correct sequence.
 - Document Delete Controls, you can delete one or more pages from the scanned images.
- Done menu list has three options selectable using the up arrow.
 - Done (Proceed with processing).
 - Hold (Suspend processing and place the Batch in the Active Existing Batches List on the Local View).
 - Cancel (End the batch, Advance it to the finished state).

Scan Procedure

- Click Scan in the Disk Scan Panel.
- Browse to and select one or more images to scan. Alternatively if you are scanning paper documents, click the view and select the ISIS or TWAIN device. If you select one of the scanner options, the Disk Scan Panel is replaced with the ISIS or TWAIN Scan pane. They have more knobs for setting Resolution, Brightness, and Contrast.
- Use the document position commands to arrange the order of the pages.
- Use the Document Delete Controls to delete unwanted images.
- Click Done to proceed with the next processing task.



Configure document, pages and fields view

- FastDoc (Admin) Datacap Server mode



- Configure the Batch Structure.
 - Add documents, pages, and fields.
- Configure Ruleset Properties.
 - Configure ruleset properties and Test profiles and rulesets.
- Configure fingerprints.
 - Add fingerprint classes, add a fingerprint page, and map page fields to image zones.

FastDoc Datacap Server mode

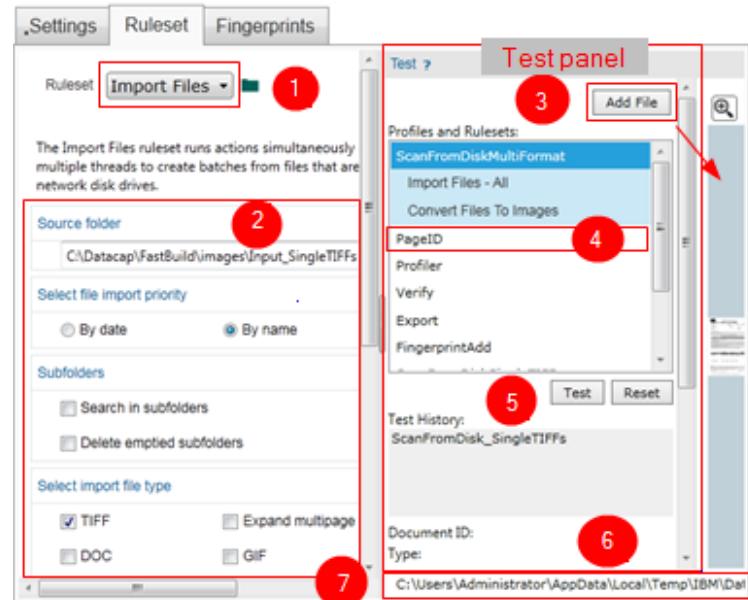
© Copyright IBM Corporation 2017

Figure 3-8. Configure document, pages and fields view



Configure Ruleset Properties

- Configure Ruleset Properties and Test profiles
 - Select a Ruleset.
 - Configure ruleset properties.
 - Add an image file.
 - Select a profile.
 - Test profiles.
 - Scroll down and check results.
 - Check error log.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-9. Configure Ruleset Properties

Error log location

C:\Users\Administrator\AppData\Local\Temp\1\IBM\Datacap\FastDoc\<application name>\TestBatch

Log files

- You enable Rulerunner error log in the Rulerunner Manager on the “Logging” tab and the “RRS Log” subtab.
- A log file is created for each job task that is run.
- When you run profiles in the FastDoc test mode, you can disable rulesets in the profile.
- If no rulesets are disabled, (That is the complete profile is run) then the errorlog file that is created is named by using the following pattern:
 - <task name>_rrs.log (Generic example).
 - convert_rrs.log (Errorlog file for the Convert task profile).
 - pageid_rrs.log (Errorlog file for the PageID task profile).
 - profiler_rrs.log (Errorlog file for the Profiler task profile).

- If one or more rulesets are disabled, then the profile task name is dropped from the error log file name.
 - _rrs.log (Used for partial task profile).

Log content

- The type of information that is captured is the progress Rulerunner records as each action in the ruleset is run.
- The type and quantity of data that is logged is set in the Rulerunner Manager on the “Logging” tab and the “RRS Log” subtab.

Example:

- Which actions are being run.
- Parameter that is passed into the action.
- Results that are returned from the action.
- Logging Application ID (Rulerunner).
- Severity level.
- Thread and Process ID.
- Time and Date stamps.
- Time difference between entries.



Configure workflow view

- Jobs.
- Profiles.
- Rulesets.

Jobs

- DemoSingleTIFFs
- Fixup Job
- Web Job
- DemoMultiFormat
- VerifyExport
- Manual Select

Profiles

Rulesets

- Convert Files To Images
- Create Documents
- Create TIFF or PDF
- Document Integrity
- Export
- Export to FileNet Content Manager
- Export to IBM Content Manager
- FingerprintAdd
- Identify Pages
- Image Enhancement
- Import Files
- Import Files - All
- Populate Fields Using Keywords
- Recognize Pages and Fields
- Routing
- Validate Fields

View Selection
Process batches of documents
Configure documents, pages, and fields
Configure workflows

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-10. Configure workflow view

Configure workflow view – has three primary areas.

Jobs - pane for selecting the defined job that you want to view or configure.

Profiles - pane for configuring the rulesets in profile. Drag rulesets into the profiles and order them.

Ruleset - selection pane. Select rulesets from here to build profiles.

- The Rulesets that are listed down the right side of the window are the rulesets available to this application.
- The list is a combination of all of:
 - The **.Rul.dll** files that are in the **C:\Datacap\<application>\dco_<application>\rules** folder.
 - The **.Rul.dll** files that are in the **C:\Datacap\RRS** folder.
- If you look at the **.rul.dll** files in these two folders, you see that most have an accompanying **.config** file. These rulesets are the UI or Compiled rulesets. Any new rulesets that you create in Datacap Studio or that are carried over from earlier Datacap versions, do not have **.config** files. Therefore, these rulesets falls into the category of a conventional ruleset.

Ruleset Types - There are now two classes of rulesets.

- Conventional rulesets - legacy or custom rulesets.
- UI Rulesets also referred to as Compiled Ruleset. Compiled Rulesets have additional information that is defined in a .config file that accompanies the rule file, which allows configuration of parameters for the ruleset within FastDoc.

Notice that in the FastDoc view the UI or Compiled rulesets are represented with three dots.

In FastDoc, the rulesets that are associated with a profile or task are listed in green boxes under the task name.



Configure Fingerprints

1. Add a Fingerprint class.
2. Add a Fingerprint image.
3. Select a Fingerprint class.
4. Select page type.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-11. Configure Fingerprints

- Fingerprints are defined to enable the identification of images by the position of data items on a page.
- When an image is identified by matching it to a fingerprint, it is quick to extract the data values from the fields that the zonal information defines that is stored in the fingerprint.
- Locating data on an image with the zonal information stored in the fingerprint is more efficient than other methods of locating data like doing key word searches.
- You can assign the fingerprints that you create to fingerprint classes to differentiate the forms that you receive from various sources.
- Fingerprint classes can be used to group fingerprints based on the source of the form.

Run Batches in Datacap Server Mode

- Start FastDoc(User).
- Click the VScan task to read documents.
 - Select image source.
- Run PagelD task and watch the activity.
 - Image Enhancement, Identify Pages, Create Documents, Document Integrity.
- Run Profiler task and watch the activity.
 - Recognize Pages and fields, Validate fields, Routing.
- Verify task.
 - Validate and Correct fields, and Submit corrected pages.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-12. Run Batches in Datacap Server Mode

- The profiles that are configured in FastDoc in the Datacap Server mode are saved on the Datacap Server.
- When Tasks are run in the Datacap Server mode, the profiles that are defined on the Datacap Server that FastDoc created, are used.
- These profiles are typically generic profiles and rulesets that are defined in the template that you used to build the application.
- The Application can be customized to any degree necessary using the Datacap Studio tool.
- Batch results are in <server>\Datacap\<Application>\batches\<date.num>.

Definitions:

- Image Enhancement – a process of correcting physical aspects of the image to make data extraction more reliable.

Examples are:

- Aligning the page so the rows are perfectly horizontal.
- Removal of horizontal or vertical lines.
- Despeckeling to remove irrelevant marks on the page.

- Inversion to convert black background to white or vice versa.
- Document Integrity – the process of checking the validity of data that is read from the image.
Two methods of integrity checking are:
 - The use of built-in algorithms to determine the degree of confidence that a character read correctly.
 - The use of validation rules for checking other field attributes like:
 - Format, string, date time, integer, float, and so on.
 - Range, for number maximum and minimum, numeric, or alphabetic.

The slide features a blue header bar with the text "IBM Training" on the left and the IBM logo on the right. Below the header is a large light blue diagonal striped background area. In the center of this area, the title "Exercise: FastDoc Datacap Server mode" is displayed in bold blue text. To the right of the title, under the heading "Requires:", there is a list of three items: "Course Exercise Guide", "Student system", and "FastDoc Datacap Server mode".

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-13. Exercise: FastDoc Datacap Server mode

Exercise objectives

- Explore FastDoc Datacap Server mode



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-14. Exercise objectives

Lesson 3.2. Create an application with Form template



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-15. Create an application with Form template

Lessons

- Introduction to FastDoc Datacap Server mode
- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-16. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- To do this operation, you must be familiar with using the FastDoc user interface in the Datacap Server mode to configure documents capture types. In this lesson, you focus on using the Form template to configure a new application.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-17. Why is this lesson important to you?

What is a FastDoc application template

- An advanced framework for creating an application.
- There are two templates released with Datacap 9.0.
 - Form.
 - Learning.
- Templates are provided to speed up application development.
- Templates can be used with Datacap Studio.
- Templates are functional applications.
- Developers can create their own templates.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-18. What is a FastDoc application template

- A template is an advanced framework for creating a new Datacap application.
- There are two templates that are released with Datacap 9.0.
 - Forms Template - For fixed form layouts like application forms, and beneficiary forms.
 - Learning Template - For more dynamic forms like invoices.
- Templates were created to speed up application development with FastDoc.
- Templates can also be used when creating an application with Datacap Studio.
- Templates are functional applications with:
 - A built-in basic document hierarchy structure.
 - A batch workflow that is made up of the basic task profiles.
 - Commonly used configurable rulesets that are already built into the task profiles.
 - A pool of more rulesets for easily expand the task profiles capability.
 - FastDoc has form design capabilities for further configuring your application.
- Save your own custom template.

- If you have a base application that you routinely use for creating new application, you can save it as a template.
- Your templates are just another application but it is good practice to never access them like an application w any of the development and production tools.
- Always use the Application Wizard to copy and rename, then access the copy.

Template Folder Structure

- Template has the same folder structure as any other application.
 - batches.
 - dco_<Template Name>.
 - export.
 - fingerprints.
 - images
- Template has the same basic internal configuration as other applications.
 - Document Hierarchy (DCO).
 - A collection of Rulesets.
 - Task profiles.
- Template definitions are in the Datacap > Templates folder.
 - Templates from the Templates folder populate the templates list to choose from in the Application Wizard.

Figure 3-19. Template Folder Structure

- Template definitions are in the Datacap > Templates folder
 - Templates that are placed in the Templates folder populate a list of templates to choose from when creating a new application with the Application Wizard.

Create an Application

- When to create a new application:
 - A new process is required.
 - A new document type.
- Before you begin.
 - Decide on an application name.
 - Decide on a template or application for a starting point.
 - Decide where to place the new application.
- Start Application Wizard.
 - From Datacap Studio.
 - From FastDoc Admin interface.
- In the Application Wizard.
 - Click Next to see the Wizard Mode options.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-20. Create an Application

When to use

The procedure to create a new FastDoc application is done whenever there is a need for a new application to process different types of documents. Also, when a capture solution is required for a new department or a document handling procedure is being implemented.

Before you begin

Know which Application or template you want to copy.

Forms Template - For fixed form layouts like application forms, and beneficiary forms.

Learning Template - For more dynamic forms like invoices.

Know what you want to name the newly created application.

Know where you want to place it.

If you are going to use it in the Datacap Server mode or mixed Datacap Server and Local mode, then the Datacap Server is the better location.



Wizard Mode window

- Select your create option:
- Create an RRS application.
 - Create an application from a template.
- Create a CMIS-based application.
 - Create a CMIS compatible application from a template.
- Copy an application.
 - Create an application from an existing application.
- Convert an 8.0.1 application to 9.0 format.
 - Convert an application that is created on Datacap 8.0.1.
- Click Next to advance.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-21. Wizard Mode window

CMIS Content Manager Interoperability Services:

- Is an open standard that CMISClient actions use to enable communication between Datacap applications and content management systems over the internet.
- This mode has the added advantage of automatically creating your document hierarchy from the document class definition on the repository.

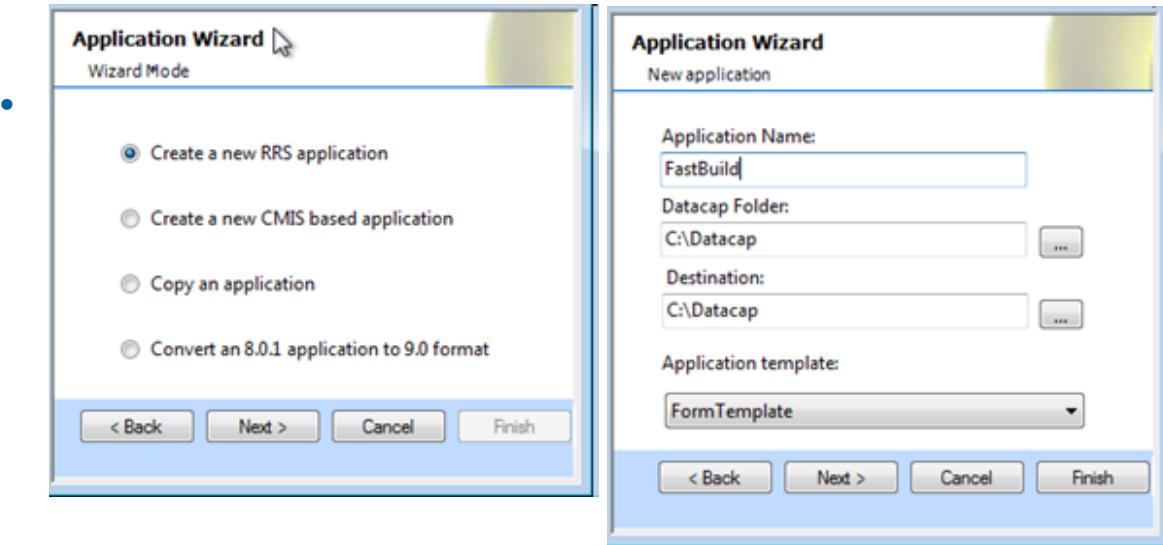
The CMISClient actions:

- Configure the connection between Datacap applications and the IBM CMIS server.
- You run these actions to:
 - Access the CMIS server.
 - Set up document attributes and folders on the server.
 - Upload documents to the server for storage.



Create a FastDoc Application

- Copy an existing application.
- Create an application from a template. (See procedure in the notes)



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-22. Create a FastDoc Application

Preserve template integrity

It is a good practice to never connect to C:\Datacap\templates\Form or Learning templates directly. Preserve the templates and create a renamed copy of the template when creating a FastDoc application.

- Start FastDoc (Admin).
- Click Datacap Application Wizard icon in the upper right.
- Click Next.
- Click Create a new RRS application.
- Click Next.

Application Name:

- Type in a name for your new application.

Datacap Folder:

- Is the root folder for locating the templates that are displayed in the application template list.
- The Datacap folder is normally on the local system but it might be on another server.
- To select a folder on another server, the path would appear as:
\<server name>\Datacap
- The wizard would expect to find a Template folder in this root folder.

Destination:

- Is the root folder for placing the application.
- This location might be on the local system but it more likely to be on the Datacap Server.
- To select a folder on another server, the path would appear as:
\<server name>\Datacap.
- The wizard would create the new application at:
\<server name>\Datacap\<application name>

Application template:

- Select a template to copy from the list. Type a name for the application that you are creating.
- For Datacap Folder browse to the Root folder on the Datacap server.
 (Network\<server name>\Datacap) or (C:\Datacap)
- For Destination, select where you want the new application to be hosted. (Network\<server name>\Datacap) or (C:\Datacap)
- Select an Application Template:
 - Form Template for known static page formats.
 - You know the types of data you want to capture.
 - You know where the data is on each image.
 - Usage examples: Application forms or Beneficiary form.
 - Learning for more dynamic documents.
 - You know the types of data you want to capture.
 - You do not know where the data is contained in the image on each image.
 - It might be different on each image.
 - Usage examples: Invoices or Shipping Inventories.
- Click Finish to complete the Application creation process.
- Click Next to advance to the Document Hierarchy and Add sample images windows.
- There is no need to configure Document Hierarchy here because it is normally done in the next phase of application creation in the FastDoc Admin Configuration interface. It can also be done in the Datacap Studio interface.
- Sample images can also be copied into the images folder later with Windows Explorer.

- Examine the Application Wizard log to verify that no errors occurred.
 - Click the Creation log link.

Next Steps

- Verify that a folder with the application name is created in the Folder that is identified in the Destination parameter.
- Start Datacap Application Manager - Start > All Programs > IBM Datacap Services > Datacap Application Manager
 - Verify that your new application appears in the menu in the left pane.
 - Select your new application and verify that database paths are created for at least the Adm, Eng, and Fingerprint database.
- **Local mode configuration**
 - While you are logged in to FastDoc (Admin) in the Local mode and if you plan to process batches in the Local mode:
 - Configure a local batch for local processing
 - Configure an upload batch for uploading the batch to the Datacap Server for further processing.
 - How to do local mode configuration was the subject of the previous three lessons.
- Run FastDoc Admin interface in the **Datacap Server – Online mode**.
- Start customizing your application to recognize, identify, and process the expected pages.

Create an application with Form template

- Application Templates.
 - Are the starting point for rapid application development.
 - Form template for structured forms.
 - Learning template for dynamic document images.
 - Custom application templates.
- Applications.
 - Any Datacap client can run them.
 - You use a template to create a FastDoc application.
 - Applications samples that are used for class exercises:
 - FastStart is a legacy Datacap 8.1 application that is built from the FastApp template.
 - FastBuild you use the Form template to develop this application.
 - TravelDocs is a sample application released with Datacap 9.0.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-23. Create an application with Form template

FastDoc Application

- A FastDoc application is just another Datacap application that any Datacap client can run. The distinguishing factor about FastDoc applications is that they use FastDoc actions that are defined specifically to interface with the FastDoc rapid application development capability.

Application Templates

- It is a good practice to keep the original copy of the application templates unmodified. When using the Datacap Application Wizard, create a new application f a template as the starting point for FastDoc Rapid Application Development. The FastDoc Rapid Application Development capability is explained later in this lesson.
 - Form template is the starting point for developing an application that processes document batches of images that have a static format with a predictable data layout like forms.
 - Learning template is the starting point for developing an application that processes documents with a more dynamic layout. That is, they require more sophisticated locate rules. An example of an application for a Learning template application is an invoice.
 - Developers can create their own custom application templates and place them in the C:\Datacap\templates folder. By placing an application in the templates folder, you enable it

to be selected from a template list when you create an application with Datacap Application Wizard.

Applications

- Any Datacap client or service can run any Datacap application, irrespective of whether it was created by using Datacap Studio, FastDoc, or both.
- An application might be referred to as a FastDoc application if it is created from one of the templates and initially configured by using the FastDoc rapid development capabilities. This development process does not prevent any client from running its tasks nor does it prevent Datacap Studio from enhancing its functions.
- FastStart - Is a FastDoc application unique to this course. It was developed with Datacap release 8.1 from the FastApp template. It was configured with FastDoc to recognize and process three different document types. FastStart was created specifically for this class and is used in Lesson 1.2. to demonstrate the processing of the three document types in the FastDoc Local mode. FastDoc Local mode is explained later in this lesson.
- FastBuild - Is a FastDoc application that you build in this course. It is created starting from the Form application template. FastBuild is created in Lesson 1.3 and configured to recognize and process two different formats of the same document type.
- TravelDocs – Is a sample application that is released with the Datacap 9.0 release. It processes car rental receipts, air ticket receipts, and hotel accommodation receipts.

Exercise: Create an application with Form template

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-24. Exercise: Create an application with Form template

Exercise objectives

- Create an application with the Form template

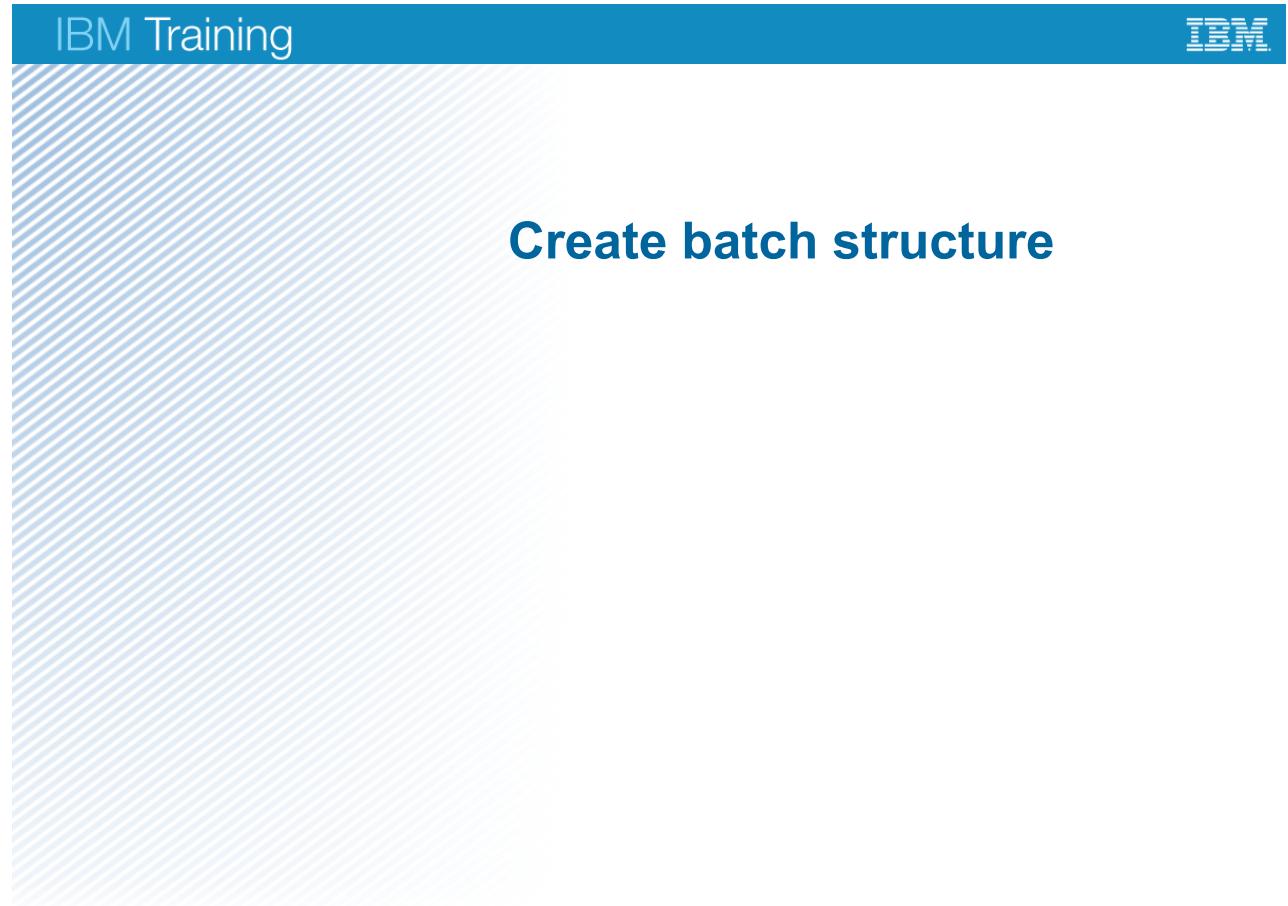


FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-25. Exercise objectives

Lesson 3.3. Create batch structure



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-26. Create batch structure

Lessons

- Introduction to FastDoc Datacap Server mode
 - Create an application with Form template
-  Create batch structure
- Configure form template rulesets
 - Define fingerprints and test rulesets
 - Populate fields with keyword lookup
 - Configure validate field rulesets
 - Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-27. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- To do this operation, you might be called upon to configure an application in FastDoc after it is created from a template with the Application Wizard. Or, you might be required to add the document recognition, and validation configuration for a document. In this lesson, you learn Configuring the Document Hierarchy for the document.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-28. Why is this lesson important to you?

Datacap Server mode

- FastDoc in the Datacap Server mode can:
 - Be used as an administrator or user interface.
 - Run as a client on a Datacap Workstation.
 - Run tasks on the Datacap server machine.
 - Configure document structure and workflows. (Admin Interface only)
- This mode is referred to as the Datacap Server Mode.
 - Also, the Online mode.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-29. Datacap Server mode



Start FastDoc and login to Datacap Server mode

- Start FastDoc (Admin) interface in the Datacap Server mode.
 - Administrator Version, FastDoc (Admin).

- Run Mode.
 - Datacap Server – Online.
 - Password = admin.

Welcome to Datacap

Local Datacap Server

Application	FastBuild
User	admin
Password	
Station	1

Login

IBM

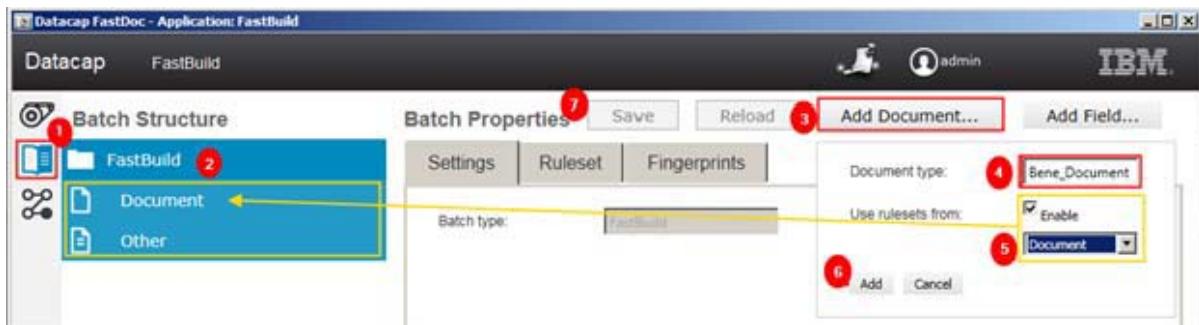
FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-30. Start FastDoc and login to Datacap Server mode



Configure a document



- Select Configure Documents, Pages, and Fields view.
- Select Batch.
- Add documents, click Add document... See Note 1.
- Name the Document type.
- Specify which ruleset group to use. See Note 2
- Complete the process, click Add.
- Save changes.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-31. Configure a document

Initiate the add documents procedure

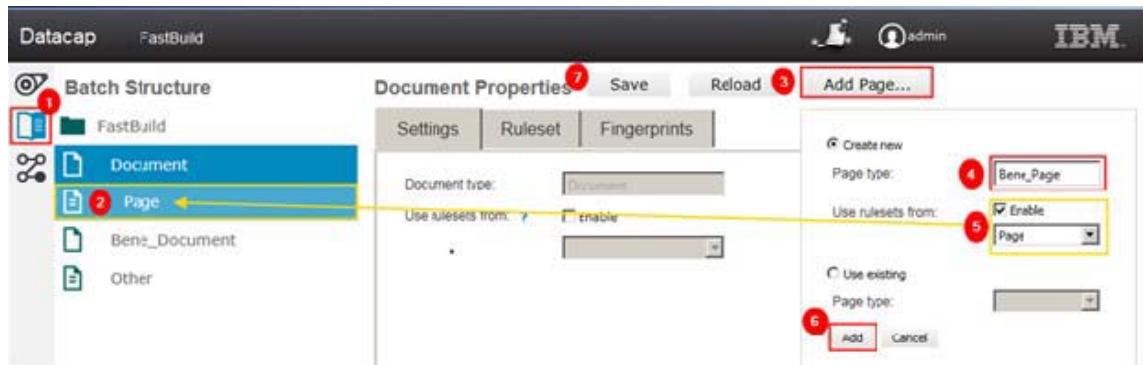
- Never edit the default document (Document).
- Always create new documents.
- Treat the Document default objects as templates that should not be changed.
- When your DCO configuration and Rulesets are customized, you can always use them as templates for other document types.

Specify which ruleset group to use

- If you don't specify, your documents are not processed as expected.
- Selecting Document in the "Use rulesets from" list causes you to use the rulesets directly out of the Form template.
- When you have defined, other documents you can select their rulesets for your new document.
- Rules are set to be shareable.

IBM Training

Configure a page



- Select Configure Documents, Pages, and Fields view.
- Select Bene_Document.
- Add page, click Add Page. See Note 1
- Name the Document type.
- Specify which ruleset group to use. See Note 2
- Complete the process, click Add.
- Save changes.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-32. Configure a page

Initiate the add documents procedure

- Never edit the default pages (Page and Other).
- Always create new pages.
- Treat the Other and Page default objects as templates that should not be changed.

Specify which ruleset group to use.

- If you don't specify, your documents are not processed as expected.
- Selecting Page in the "Use rulesets from" list causes you to use the rulesets directly out of the Form template.
- When you defined other pages, you can select their rulesets for your new document.



Document integrity

- Verify “Use Rulesets from” Document.
 - “Use ruleset from” is enabled and Document is selected.
- Verify “Use Rulesets from” Page.
 - “Use ruleset from” is enabled and Page is selected.
 - Set Minimum, Maximum, and Order all to 1.
- Save.

The screenshot shows the IBM FastDoc Datacap Server mode interface. On the left, the 'Batch Structure' panel displays a tree view with nodes: 'FastBuild', 'Document', 'Bene_Document', 'Bene_Page' (which is selected and highlighted in blue), and 'Other'. On the right, the 'Page Properties' panel has tabs for 'Settings', 'Ruleset', and 'Fingerprints'. The 'Settings' tab is active. It shows 'Page type: Bene_Page', 'Use rulesets from: ?' with a checked 'Enable' checkbox, and a dropdown set to 'Page'. A note below explains the settings for determining page types in a document. Under 'Minimum', 'Maximum', and 'Order', each has a value of '1'.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-33. Document integrity

Verify Use Rulesets from Document

- On the Bene-Document object in the Batch Structure on the Settings tab, verify that “User ruleset from” is enabled and Document is selected.

Verify Use Rulesets from Document

- On the Bene-Page object in the Batch Structure on the Settings tab, verify that “User ruleset from” is enabled and page is selected.

Set Minimum, maximum, and Order

- These properties are:
 - Not inherited from page rulesets.
 - Required otherwise the document is not created for each page.

Remember to always save your changes.



Add Fields

You can add fields at the Batch, Document, and Page objects.

The screenshot shows the IBM Datacap interface with the following details:

- Batch Structure:** A tree view on the left with nodes: FastBuild, Document, Bene_Document, Bene_Page (selected), and Other.
- Page Properties:** A dialog box titled "Page Properties" with tabs: Settings (selected), Ruleset, Fingerprints. Buttons: Save (with a red circle 5), Reload, Add Field... (with a red circle 2).
- Settings Tab:**
 - Page type: Bene_Page
 - Use rulesets from: Page (checkbox checked)
 - Enable (checkbox checked)
- Create new section:**
 - Field type: Employee_Name (with a red circle 3)
 - Use rulesets from: Enable (checkbox checked)
- Use existing section:**
 - Field type: (dropdown menu)
- Buttons:** Add (with a red circle 4) and Cancel.

- Select Object in the Batch Structure.
- Start the add field procedure, click Add field ...
- Name the field.
- Complete the process, click Add.
- Save changes.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-34. Add Fields

No fields are defined in the form template so there are no field Rulesets to inherit.

Field Settings

- Visible option must always be set
- Other options:
 - Display name: is the label on the Verify panel.
 - Use rulesets from: Enable. See note
 - SQL select: for getting selection list components from a database.
 - SQL lookup: populate a field with the selected value.
 - Data type mask: Select the type of value. See notes
 - Optical mark: Enable
 - Dictionary: Add an entry for each optical mark.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-35. Field Settings

Visible option

- If Visible option is not set, then the field does not appear on the verify panel.

Use rulesets from:

- On the previous create field slide, this option was not selected because there were no fields that are defined in the Forms template. If you are defining a new page, and a field of the same type is already configured with the same processing requirement, you can select it here.

Data type:

- Alphanumeric
- Integer
- Floating Point
- Date
- Time hh:mm:ss
- Currency

Context help: Click the? mark to display context-sensitive help for those options.

Exercise: Configure batch structure

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-36. Exercise: Configure batch structure

Exercise objectives

- Create a batch structure

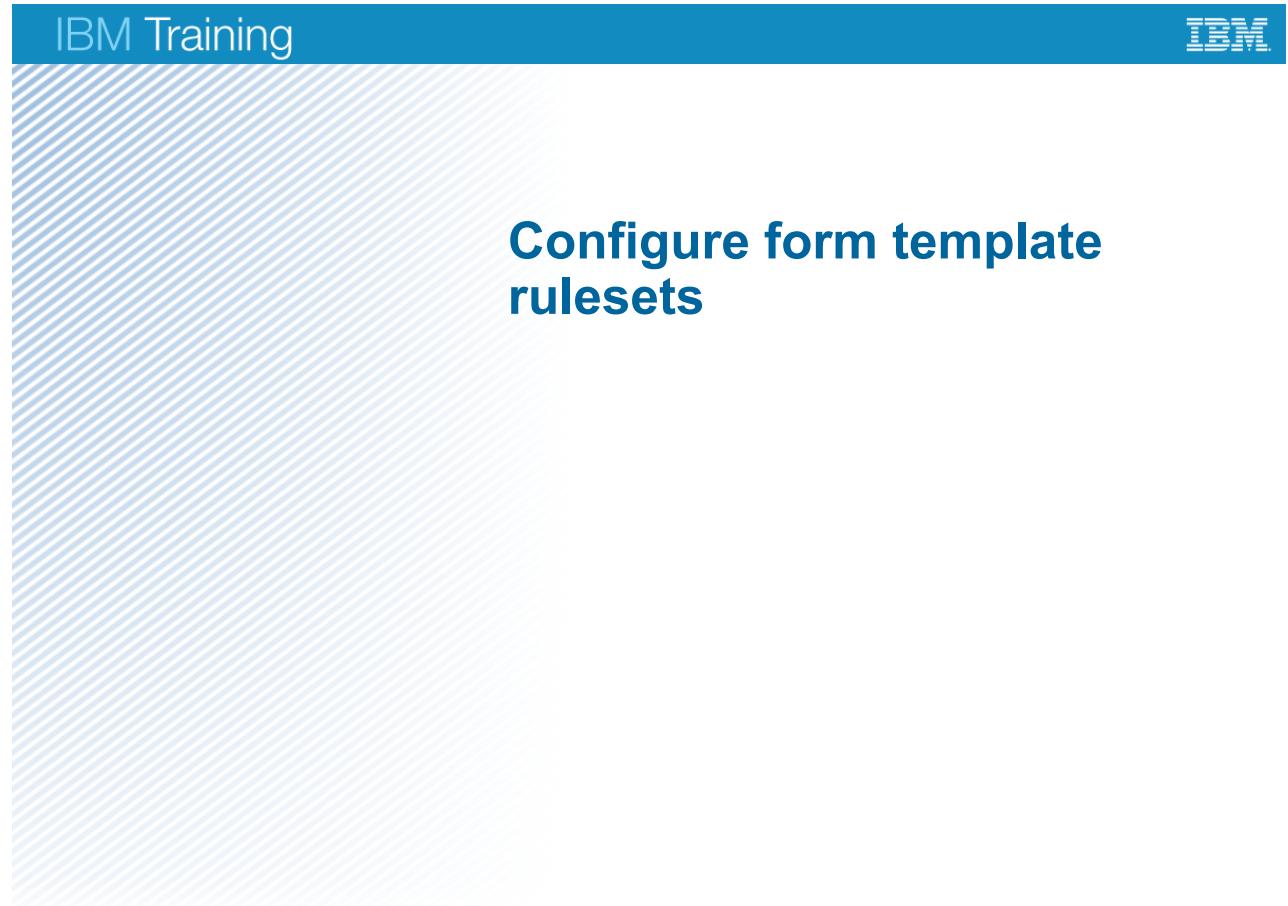


FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-37. Exercise objectives

Lesson 3.4. Configure form template rulesets



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-38. Configure form template rulesets

Lessons

- Introduction to FastDoc Datacap Server mode
- Create an application with Form template
- Create batch structure
-  Configure form template rulesets
 - Define fingerprints and test rulesets
 - Populate fields with keyword lookup
 - Configure validate field rulesets
 - Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-39. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- You are required to configure an application in FastDoc after it is created from a template with the Application Wizard. In this lesson, you learn to configure the page identification and field recognition rulesets.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-40. Why is this lesson important to you?

Configure Forms template rulesets

- Batch level.
 - Import Files.
 - Identify Pages.
 - Convert Files To Images.
- Create Documents.
- Document level.
 - None.
- Page level.
 - Image Enhancement.
 - Recognize Pages and Fields.
- Field level.
 - Recognize Pages and Fields.
 - Validate Fields.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-41. Configure Forms template rulesets

Configuring rulesets for the application

- After you set up the documents and fields to process, you can configure the rulesets to run on the application that is based on the template that you are using.
- You configure rulesets to run on the batches, documents, pages, and fields that are in the application in Datacap Server mode.

Batch level rulesets

- Import Files - no configuration is required for Form template.
- Identify Pages – no configuration is required for Form template.

Configure how Page Identification is done.

Default settings:

- Fingerprinting.
- Convert files To Images.

Convert file images to a different format before processing. This ruleset is only needed if the import files are other than TIFF.

- Create Documents.

Default settings: No Options set.

Document level rulesets

No Document Rulesets.

Page level rulesets

- Image Enhancement.

Default settings:

- Deskew.
- Remove Lines.
- Border Removal.

- Recognize Pages and Fields.

Default settings: No Options set.

Field level rulesets

- Recognize Pages and Fields.

Default settings: No Options set.

- Validate Fields.

Default settings: No options set.

Configure rulesets general procedure:

1. Start FastDoc in Datacap Server mode.
2. On the Datacap window, click the Configure workflow icon.
3. Select the job that you want to run and click the Configure documents, pages, and fields icon.
4. In the Batch Structure pane, select the name of the application and click the Rulesets tab.
5. Configure the following rulesets for the appropriate batch structure level in the application:
6. Use the Test tab to confirm that the processing is working correctly and make corrections as needed.
7. Process real documents.

Import files - Batch level ruleset

- Create batches from files that are stored on a local or network drives.
- Import Files UI properties.
 - Source folder.
 - Example: C:\Datacap\<application name> images\Input_SingleTIFFs
 - Priority by date or alphabetically by name.
 - Subfolders.
 - Search in subfolders.
 - Delete emptied subfolders.
 - Select import file type.
 - Maximum number of files to import.
 - Copy image folder.
 - Problem image folder.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-42. Import files - Batch level ruleset

Import Files - no configuration required.

- Default settings are:
- Source and Copy folder = C:\Datacap\<application name> images\Input_SingleTIFFs.
- Select file import priority = By name.
- Import File type = TIFF.
- Maximum number of files to import = 50.

Import Files UI properties

- Source folder.
- Define the folder to read input files from.
- Priority by date or alphabetically by name.
- By date - causes documents in the batch to be sorted by date.
- By name - causes documents in the batch to be sorted by name.
- Subfolders.
- Search in subfolders.

Also, search the subfolders of the Source folder for more images to scan.

- Delete emptied subfolders.

When images are read and relocated delete remaining empty subfolders.

- Select import file type

Select the file types to expect in the source files.

If more than one file type is expected, then you can use the Convert Files To Images ruleset to convert them all to .TIFF format before processing.

- Maximum number of files to import.

This option allows large volumes of pages to be read in manageable size batches.

Options are: 1, 10, 25, 50, 100, All or set some custom value.

- Copy image folder

Images are copied to this folder when the batches are queued.

- Problem image folder.

- Any unidentifiable pages are placed here.

Convert Files To Images – Batch level ruleset

- Convert file images to a different format before processing.
- This ruleset is only needed if the import files are other than TIFF.
- Convert Files To Images UI properties.
 - File naming pattern.
 - Hierarchical – name = ADADADAD.xxx.
 - Sequential – name = TMnnnnnnn.xxx.
 - Source Page.
 - Remove page after processing.
 - Set page to DELETED after processing.
 - Leave Pages in batch after processing.
 - Only process pages with status scan OK.
 - File formats to convert.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-43. Convert Files To Images – Batch level ruleset

Convert Files To Images

- File naming pattern = Sequential.
- Remove pages after processing.
- All file formats selected for conversion.

Convert Files To Images ruleset UI properties

- File formats to convert
- Bitmap
- Multi-page TIFF
- ZIP Archive Expansion
- PDF
- Microsoft Word Conversion
- Microsoft Excel Conversion
- HTML Conversion
- RTF Conversion

Image enhancement – Batch level ruleset

- Goal of image enhancement.
 - Eliminate artifacts that interfere with recognition.
- Affects of image enhancement.
 - Option that you select affect all images that the application processes.
- When to complete image enhancement.
 - On fingerprint images, when you are setting up the fingerprint library.
 - Again on your document images after input but before page identification.
- Configure image enhancement.
 - In Batch Structure, select page.
 - Click the Ruleset tab and select Image Enhancement ruleset.
 - Open image file...
 - Configure options and view results immediately in the dual image pane.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-44. Image enhancement – Batch level ruleset

Goal of image enhancement

- To eliminate lines, shading, misalignment, and other artifacts that can interfere with the recognition process.

Affects of image enhancement

- Image enhancement is typically done before the page type is known, in other words, before page identification. You must set up the image-processing properties in a way that works well for all page types.

When to complete image enhancement

- You complete image enhancement on fingerprint images when you are setting up the fingerprint library.
- You must complete it again on your document images after input but before page identification.

Configure image enhancement

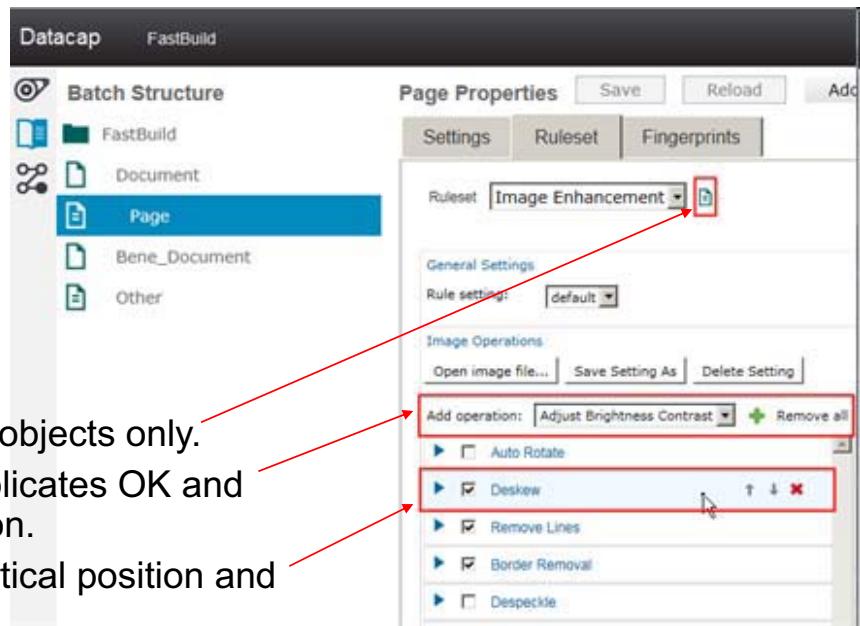
- Default options that are inherited from the Form template Page object are:
- Deskew
- Remove Lines
- Boarder Removal

IBM Training



Image Enhancement ruleset

- Applies to Page objects only.
- Add options Duplicates OK and Remove all option.
- Hover to see vertical position and remove controls.



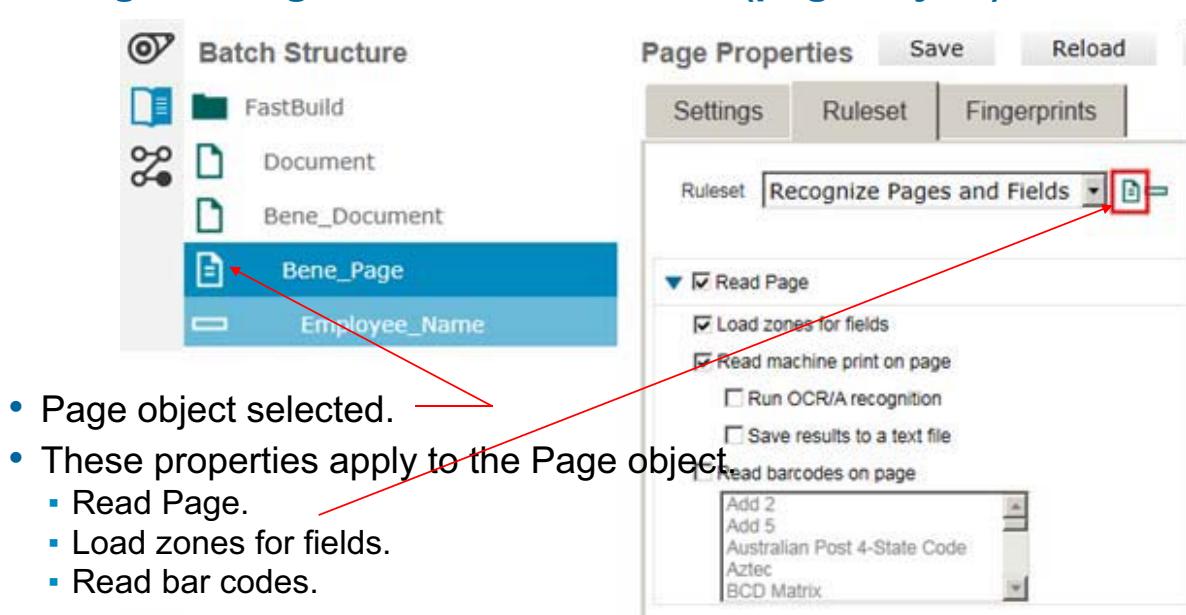
FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-45. Image Enhancement ruleset

IBM Training 

Recognize Pages and Fields ruleset (page object)



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-46. Recognize Pages and Fields ruleset (page object)

Page Properties

- Read Page
- You need to select read the page to read all of the page content. The resulting finger print is compared against saved fingerprints to identify the page.
- Load Zones for field.
- Load the defined field zones from the fingerprint file.
- Read machine print on page.
- Select the option for reading machine print.
 - Run OCR/a recognition.
 - Save Results to a text file.
- Don't use this option if your fields are handwritten. It is very time consuming to continually search for machine code with many hand written fields.

Place a text file in the batches folder that contains all of the text that is read from the page. It is the text version of the fingerprint.

- Read bar code on page (Select this option if you identify the page with a bar code).



Fields properties settings

- Field type.
- Display name.
- Use rulesets from.
- Visible.
- SQL select.
- SQL lookup.
- Data type mask.
- Optical mask.
- Allow multiple settings.

Field Properties Save Reload Add Field...

	Settings	Ruleset	Fingerprints
Field type:	State		
Display name:	.		
Use rulesets from:	<input type="checkbox"/> Enable		
Visible:	<input checked="" type="checkbox"/>		
SQL select:			
SQL lookup:			
Data type mask:	Not Specified		
Optical mark:	<input type="checkbox"/> Enable		
Dictionary:		<input type="checkbox"/> Allow multiple selections	
	Value	Display Text	
	Add		

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-47. Fields properties settings

Field type

- Name that is assigned when the field was created.

Display name

- An optional name to substitute for the Field type when the field is displayed.

Use rulesets from Enable

- Inherit rulesets for this field from another field.

Visible

- It is important that this field is always set or the field does not appear on the verify panel.

SQL select

- A SQL statement to extract the required value from a SQL database.

SQL lookup

- A SQL statement to populate the field with a lookup on a SQL database.

Data type mask

- A list to select a specific data type:

- Alphanumeric.
- Integer.
- Floating Point.
- Date.
- Time HH:MM:SS.
- Currency.

Optical mask Enable

- Define this field as an Optical Mask Recognition field.
- See next slide for more detail.

Allow multiple settings Enable

- See next slide.

Dictionary

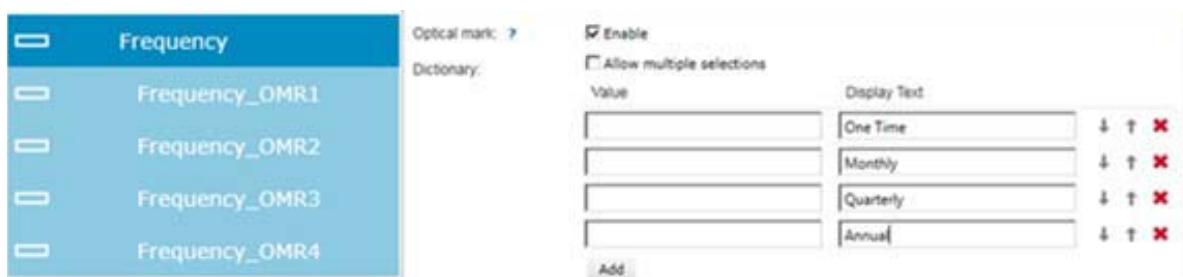
- See next slide.

ADD

- See next slide.

IBM Training 

Fields properties settings



- Optical mask Enable.
- Allow multiple settings.
- Dictionary.
- Add.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-48. Fields properties settings

Data type mask

- A list to select a specific data type:
- Alphanumeric
- Integer
- Floating Point
- Date
- Time HH:MM:SS
- Currency

Optical mask Enable

- Define this field as an Optical Mask Recognition field.
- The field consists of one or more:
- Check boxes
- Radio buttons
- Other mark objects

Allow multiple settings Enable

- If the field is an OMR field, allow multiple options to be checked or selected.

Dictionary

- Define optical marks fields in the OMR group

ADD

- Option to initiate adding OMR items to the dictionary.
- As items are added to the dictionary, the mark object is added to the mark field group.
- Frequency_OMR1 ? One Time
- Frequency_OMR2 ? Monthly
- Frequency_OMR3 ? Quarterly
- Frequency_OMR4 ? Annual

IBM Training 

Recognize Pages and Fields ruleset (field object)

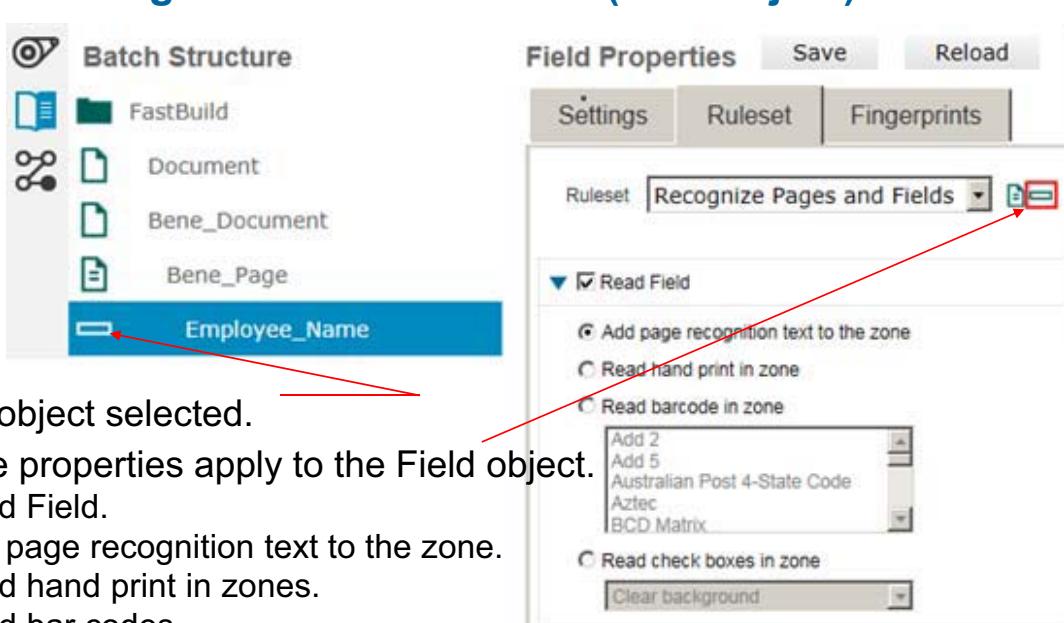


Figure 3-49. Recognize Pages and Fields ruleset (field object)

Field recognition methods are:

- Zonal.

Most reliable and most widely used for forms applications.

- Key word search.

Used the position of a value can be at different places on the image.

Example: the total of a list of purchased item. Position varies depending on the length of the list.

- Regular Expression.

Used to recognize fields with a specific format.

Examples: Social security number xxx-xx-xxxx or a 10 digit phone number xxx-xxx-xxxx

Field Properties

- Read Field

Read field is always required for zonal recognition.

- Add page recognition text to the zone.

Populate the zone with fields read during page recognition.

- Read hand print in zones.
Select this option if the zone has hand printed data.
- Read bar code in zones.
Select this option if the zone has a bar code.
- Read check boxes in zone.
Select this option if the zone has a check boxes.

Zonal recognition engines that are currently used are Abby or ICRC.

Exercise: Configure form template rulesets

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-50. Exercise: Configure form template rulesets

Exercise objectives

- Configure form template rulesets



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-51. Exercise objectives

Lesson 3.5. Define fingerprints and test rulesets

Define fingerprints and test rulesets

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-52. Define fingerprints and test rulesets

Lessons

- Introduction to FastDoc Datacap Server mode
 - Create an application with Form template
 - Create batch structure
 - Configure form template rulesets
-  Define fingerprints and test rulesets
- Populate fields with keyword lookup
 - Configure validate field rulesets
 - Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-53. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- You are required to configure an application in FastDoc after it is created from a template with the Application Wizard. In this lesson, you learn to configure a fingerprint for a page image.

FastDoc Datacap Server mode

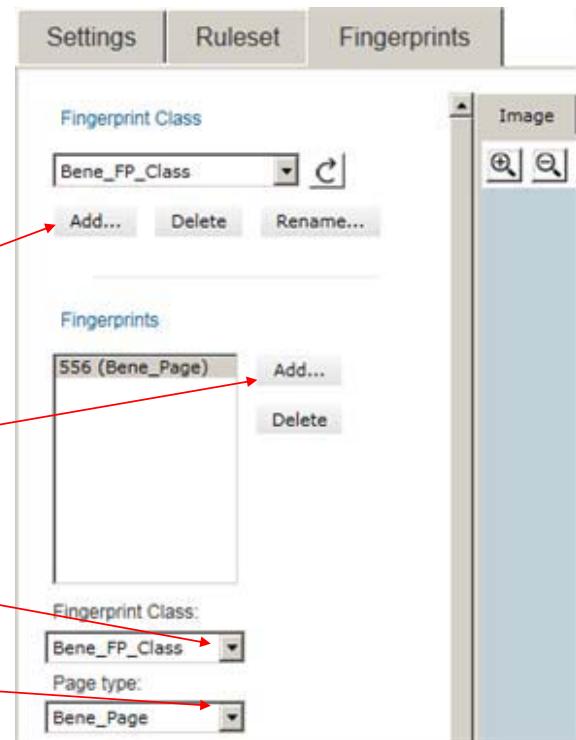
© Copyright IBM Corporation 2017

Figure 3-54. Why is this lesson important to you?

IBM Training

Fingerprint - Define Fingerprint

- Add Class.
- Add Fingerprint.
- Select a fingerprint class.
- Select a fingerprint page type.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-55. Fingerprint - Define Fingerprint

To configure a Fingerprint for your image

1. Optionally define a fingerprint class with the Fingerprint Class > Add....
If you don't define a class, your fingerprint is added to the New class.
2. Add a fingerprint image to the image pane with Fingerprint > Add.
3. Select a fingerprint class.
4. Select a fingerprint page type.
5. Select each field in the batch structure for the selected page type and use click'n'key to identify the field zones on the image.

Warning: Use Datacap Studio to define check boxes.



Fingerprint – Select Zones

I
Safe & Secure

BENEFICIARY DESIGNATION FORM

INSTRUCTIONS: PLEASE PRINT, SIGN AND DATE THIS FORM IN BLACK INK.

Employee Name	Mary Smith	Social Security No	245-67-8901	Date of Birth	05-11-1966	Home Telephone Number	555-234-5678
Home Address	2345 Maple Street	City	Smarter City	State	NY	Zip Code	10026
Employer	Company A						Group Number
							A12345678

Irrevocable Beneficiary: Yes No X

Zone

Name: EmployeeName

Position:

Left:	206	Top:	375
Right:	1010	Bottom:	442

Clear zone

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-56. Fingerprint – Select Zones

To configure a Fingerprint for your image

1. Optionally define a fingerprint class with the Fingerprint Class > Add....
If you don't define a class, your fingerprint is added to the New class.
2. Add a fingerprint image to the image pane with Fingerprint > Add.
3. Select a fingerprint class.
4. Select a fingerprint page type.
5. Select each field in the batch structure for the selected page type and use click'n'key to identify the field zones on the image.

Warning: Use Datacap Studio to define check boxes.

What is design time testing?

- FastDoc's built in ability to test:
 - The page identification to identify the page and document types.
 - Page and field recognition to recognize and extract the data.
- Test Steps (see test procedure in the notes)
 - Read an image.
 - Run the PageID task profile.
 - Run the Profiler task profile.
- Verify results
 - In the Test pane.
 - In the Test batches folder.
- Test batch folder path at the bottom of the Test pane.
 - C:\Users\Administrator\AppData\Local\Temp\1\IBM\Datacap\FastDoc\FastBuild\TestBatch
 - Click Open.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-57. What is design time testing?

FastDoc Built-in test interface

FastDoc immediately runs the PageID and Profiler task to verify that the Fingerprint is correctly defined and the rulesets are correctly configured.

Test procedure

1. Click the Ruleset tab.
2. Add a file c:\DC9-Lab Exervices\FastBuild\Input_SingleTIFFsBene001.tif (Equivalent to VScan step).
3. Select the PageID profile in the test panel Profiles and Rulesets pane. (PageID is the next task after VScan).
4. Verify results:
 - Document type is Bene_Document.
 - Page type is Bene_Page.
 - Fingerprint is 556.
5. Select the Profile profile.
6. Click the Routing ruleset to disable it.

7. Click Test.

8. Verify the results:

- Click each field in the Batch Structure pane on the Bene_Page page.
- Verify that the value is correctly interpreted for each.
- Verify that the status is Pass for each field.



IBM

Test pane profiles and rules

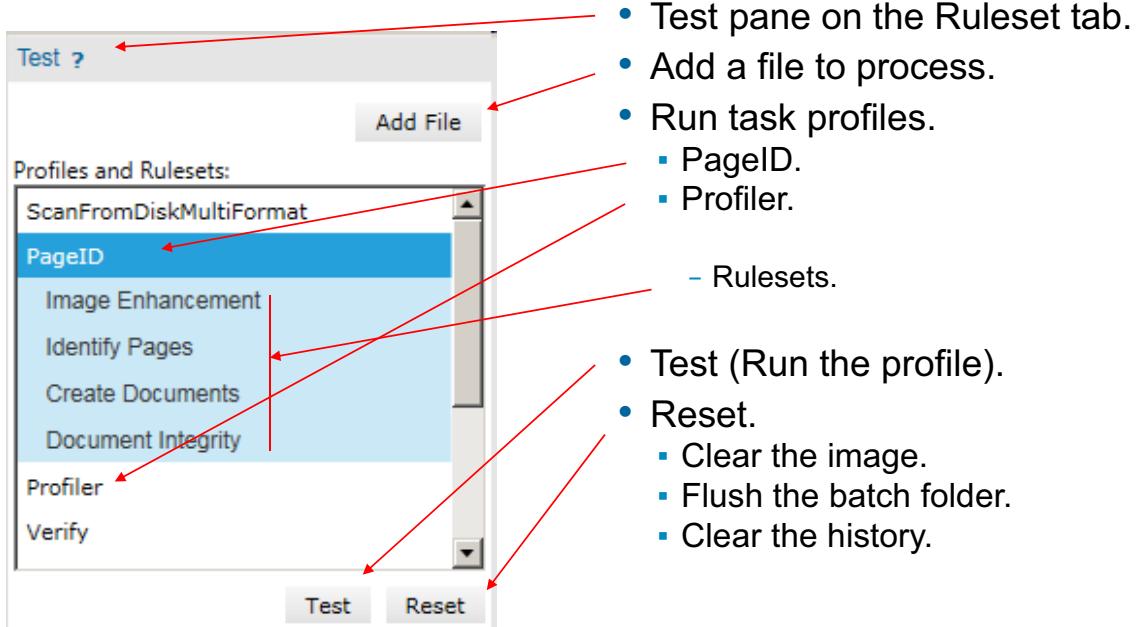
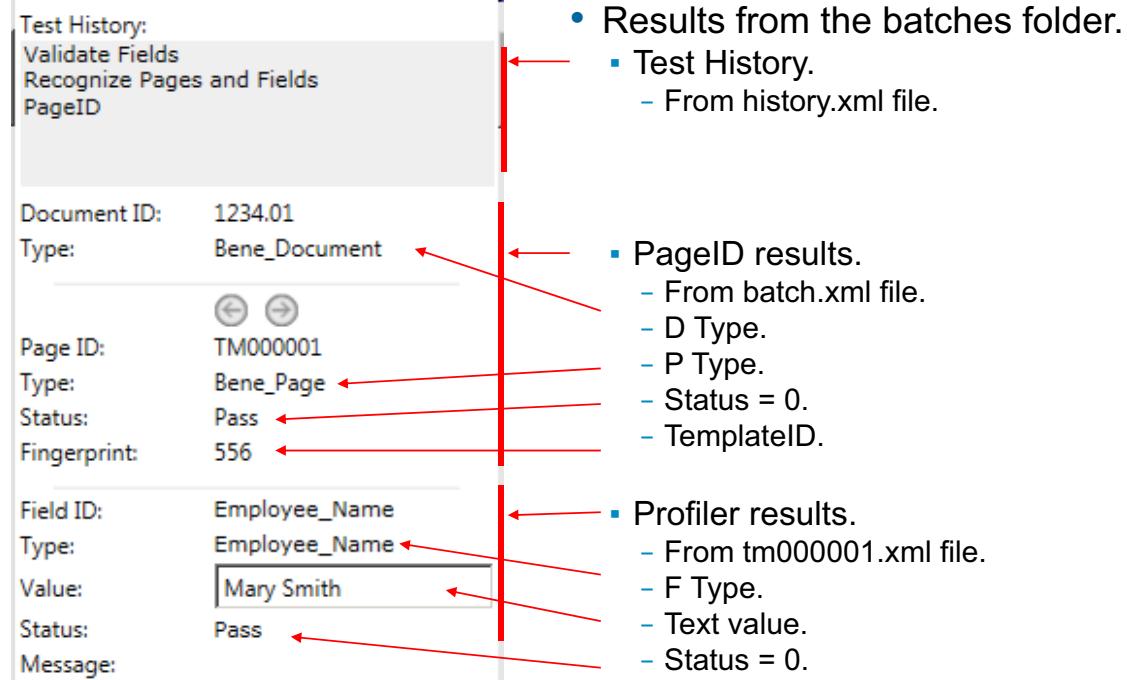


Figure 3-58. Test pane profiles and rules



Test page history and results



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-59. Test page history and results

history.xml

```
<History>
  <Item>Validate Fields</Item>
  <Item>Recognize Pages and Fields </Item>
  <Item>PageID</Item>
</History>
```

batch.xml

```
<?xmlstylesheet type="text/xsl" href=".\\..\\dco.xsl"?>
<B id="1234">
  <V n="STATUS">0</V>
  <V n="TYPE">FastBuild</V>
  <D id="1234.01">
    <V n="TYPE">Bene_Document</V>
    <V n="STATUS">0</V>
```

```

<P id="TM000001">
    <V n="TYPE">Bene_Page</V>
    <V n="STATUS">0</V>
    <V n="IMAGEFILE">tm000001.tif</V>
    <V n="processingSettings"></V>
    <V n="OriginalImage">
        c:\users\administrator\appdata\local\temp\1\ibm\datacap\fastdoc\fastbuild\testbatch\
        TM000001.tio</V>
    <V n="s_srCreateCCO">1</V>
    <V n="UseOutOfProcessRecog">TRUE</V>
    <V n="Confidence">1</V>
    <V n="Image_Offset">0,0</V>
    <V n="TemplateID">556</V>
    <V n="Fingerprint Created">No</V>
    <V n="DATAFILE">tm000001.xml</V>
    <V n="RecogStatus">0</V>
    <V n="CCOFILE">
        c:\users\administrator\appdata\local\temp\1\ibm\datacap\fastdoc\fastbuild\testbatch\
        tm000001c.xml</V>
    <V n="ZoneRead">FPXML</V>
    <V n="ErrorMessage"></V>
</P>
</D>
</B>

tm000001.xml
<?xmlstylesheet type="text/xsl" href=".\\.\\dco.xsl"?>
<P id="TM000001">
    <F id="Employee_Name">
        <V n="TYPE">Employee_Name</V>
        <V n="Position">206,375,1010,442</V>
        <V n="STATUS">0</V>
        <C cn="10" cr="216,395,247,425">77</C>
        <C cn="10" cr="250,402,272,425">97</C>
        <C cn="10" cr="273,402,287,425">114</C>
        <C cn="10" cr="287,403,307,434">121</C>
    </F>
</P>

```

```
<C cn="10" cr="424,394,425,426">32</C>
<C cn="10" cr="321,394,344,426">83</C>
<C cn="10" cr="348,402,378,425">109</C>
<C cn="10" cr="384,395,389,425">105</C>
<C cn="10" cr="391,396,402,425">116</C>
<C cn="10" cr="404,395,424,425">104</C>
</F>
```

Test pane error resolution

- PageID task.
 - Document and page not correctly identified.
 - Check the fingerprint tab. Make sure Fingerprint 556 is set to Bene_Page type and not the Other type.
- Profiler task.
 - Field value was not read.
 - Make sure the PageID successfully identified the page as a Bene_Page.
 - Check your RuleSet tab and make sure that the settings are correct on the Recognize Pages and Fields ruleset for the Bene_Page and each of the fields.
- Still have a problem?
 - Look at the log files in the batches folder.
 - pageid_rrs.log and _rrs.log.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-60. Test pane error resolution

Still have a problem?

- TestBatch folder is at:

C:\Users\Administrator\AppData\Local\Temp\1\IBM\Datacap\FastDoc\FastBuild\TestBatch

Note: At the bottom of the test pane, click Open to get to the TestBatch folder

Exercise: Define fingerprints and test rulesets

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-61. Exercise: Define fingerprints and test rulesets

Exercise objectives

- Add a page fingerprint
- Design time testing of rulesets



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-62. Exercise objectives

Lesson 3.6. Populate fields with keyword lookup



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-63. Populate fields with keyword lookup

Lessons

- Introduction to FastDoc Datacap Server mode
- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
-  Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-64. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- You are required to configure an application in FastDoc after it is created from a template with the Application Wizard. In this lesson, you learn to recognize the field and extract the data with the keyword lookup method rather than the zonal recognition.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-65. Why is this lesson important to you?



Add the Locate Fields Using Keyword ruleset

- On the FastDoc Configure workflow view.
- Add the Populate Fields Using Keywords ruleset to the Profiler task profile.



The screenshot shows the FastDoc Configure workflow view. On the left, there is a sidebar with a green header labeled "Profiler". Below it is a list of tasks: "Recognize Pages and...", "Populate Fields Using...", "Validate Fields", and "Routing". A red arrow points from the "Populate Fields Using..." button towards the "Rulesets" column on the right. The "Rulesets" column lists various global rulesets, including "Convert Files To Images", "Create Documents", "Create TIFF or PDF", "Document Integrity", "Export", "Export to FileNet Content Manager*", "Export to IBM Content Manager", "FingerprintAdd", "Identify Pages", "Image Enhancement", "Import Files", "Import Files - All", and "Populate Fields Using Keywords". The "Populate Fields Using Keywords" rule is highlighted with a blue background. At the bottom of the screen, there is a copyright notice: "© Copyright IBM Corporation 2017".

Figure 3-66. Add the Locate Fields Using Keyword ruleset

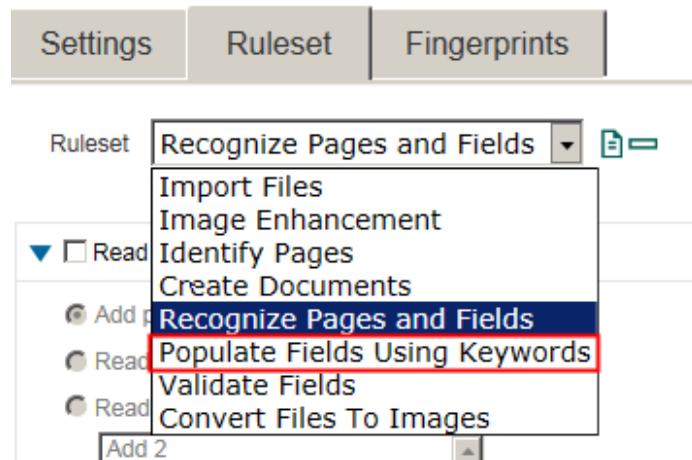
Add the Populate Fields Using Keywords ruleset:

- The Rulesets column on the right pane contains all of the global rulesets from the Datacap\RRS folder.
- Drag the “Populate Field Using Keywords” ruleset from the Rulesets column to the Profiler task profile.
- Place the “Populate Field Using Keywords” ruleset after the “Recognize Pages and Fields” ruleset.
- Use the Save control to save your changes.



Configure Populate Fields Using Keyword ruleset

- On the Configure documents, pages, and fields view.
- Select the Field in the Batch Structure pane.
- On the Ruleset tab, select the Populate Fields Using Keywords.



FastDoc Datacap Server mode

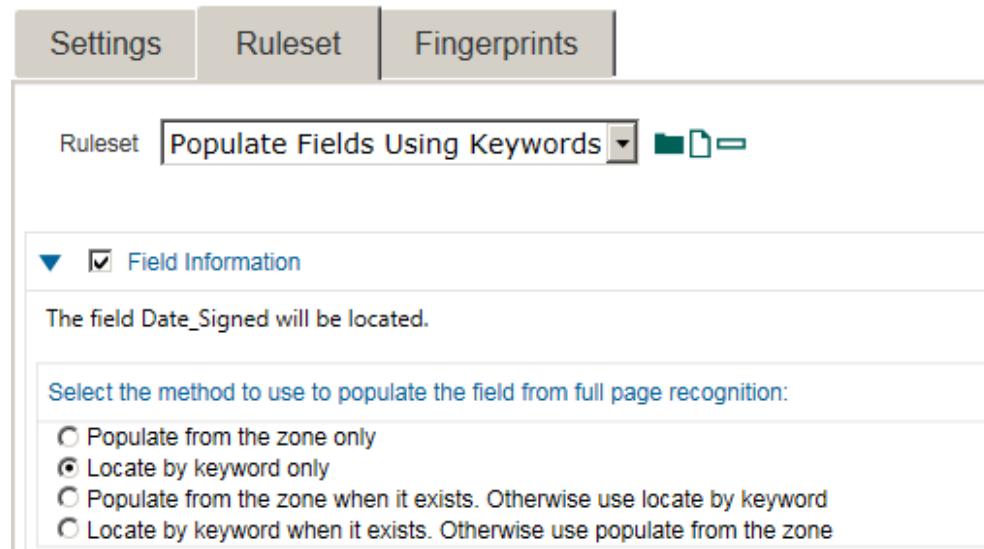
© Copyright IBM Corporation 2017

Figure 3-67. Configure Populate Fields Using Keyword ruleset

IBM Training



Select population methods



- Always set the Field information check box.
- Select populate option.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-68. Select population methods

- Always set the Field information check box.
 - If not set then no field population is done.
- Population options are:
 - Use zone.
 - Use keyword search.
 - Use a prioritized combination of both options to ensure reliable field recognition.

IBM Training

Locate the lookup term

- Sample search label and value.
Date Signed 08-31-2012
- Click Add Term.
 - Type the label to search for.
- Select search options.

Select the search type:

Locate terms allowing for common recognition errors
 Locate terms using regular expression syntax

Select to locate the first occurrence or last occurrence on the page:

Locate first occurrence
 Locate last occurrence

Select the order of the search terms:

Search for terms in sequential order
 Search for all of the terms at the same time

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-69. Locate the lookup term

Add Term

- You can add more than one term to search for. When you are searching for a total value, possible key words to search for might be:

Examples:

Total

Total:

Total Sale

Total Sale:

Sum

If any of these terms are found on the page, then the subsequent directions are followed to locate the value.

Select Search options

- Select the search type.
 - Locate terms and allow for common recognition errors.

This option does a normal test search.

- Locate terms by using regular expressions syntax.

This option uses regular expression to search for patterns or specific characters or character ranges.

- Select to locate the first occurrence or last occurrence on the page:

- Locate first occurrence.

This option results in a search from the top of the page to locate the first occurrence.

- Locate last occurrence.

This option results in a search from the bottom of the page to locate the first occurrence.

- Select the order of the search terms:

- Search for the terms sequentially

This option searches the entire document for each defined term sequentially.

- Search for all of the terms at the same time.

This option checks each term for each test field on the page and ends when the first term is found.



Locate value with direction and grouping

- Select direction options
- Select grouping options
 - See notes for option descriptions

Specify movement to the desired word starting from the located word:

Direction	Count
<input type="button" value="Move to next word"/> ... <input type="button" value="Add Locate Movement"/>	<input type="button" value="X"/>

Specify grouping of words to include in the found field:

Group direction:	<input type="button" value="No grouping"/> ...
Group spacing count:	

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-70. Locate value with direction and grouping

Add Locate Movement

- Direction options.
 - Move down a line.
 - Move to next word.
 - Move to previous word.
 - Move up a line.
- Direction count.
 - Specify the number of lines to move up or down or the number of words to move left or right.
- Group directions options.
 - Group leading and trailing word.
 - Group leading words.
 - Group trailing words.
 - No grouping.

Examples:

For the Employee name, you select "group trailing words" to make sure that the employee first and surnames were read.

Therefore, for the Date Signed field no grouping is required because the date consists of contiguous characters, 08-31-2012 and interpreted as one word.

- Group spacing count.
 - The number of words that separate the words that should be considered the same word.

Exercise: Populate fields with keyword lookup

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-71. *Exercise: Populate fields with keyword lookup*

Exercise objectives

- Populate fields with keyword lookup



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-72. Exercise objectives

Lesson 3.7. Configure validate field rulesets

Configure validate field rulesets

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-73. Configure validate field rulesets

Lessons

- Introduction to FastDoc Datacap Server mode
- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
-  Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-74. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system. The preferred way of creating new application is with FastDoc rapid application builder capabilities.
- You are required to configure an application in FastDoc after it is created from a template with the Application Wizard. In this lesson, you learn to configure validation rules to establish the validity of fields that are extracted from the image.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-75. Why is this lesson important to you?

Validate Fields ruleset

- Validate Selected fields.
- Validate this field, <field name>.
 - Must be set for any validation to take place.
- Skip validation if the field is empty.
- Each enabled validation section is evaluated in order.
- Validation check groups.
 - Common.
 - Length.
 - Value Content.
 - Date.
 - Currency and numeric.
 - Lookup.
 - Custom Error Message.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-76. Validate Fields ruleset

The Validate Fields ruleset is a compiled (UI) ruleset that provides a wide range of field validation options for the selected value. If the required degree of validation cannot be accomplished with these validation options, then you must use Datacap Studio to create a conventional ruleset from the validate action libraries. The use of Datacap for creating specialized rulesets is covered in a later class.

Common ruleset options

- Common.
 - Overridable. (**default** option in Form template)
 - Trim spaces.
 - Covert characters options:
 - No change (**default** option in Form template)
 - To lowercase.
 - Convert to uppercase.
 - Allow only characters.
 - Filter strings.
 - Replace.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-77. Common ruleset options

Common

Sets up the validation characteristics for the captured values of the current field.

- Overridable. (**default** option in Form template)

The verify operator can override the recognized value.
- Trim spaces

Delete spaces at the beginning and end of the captures value.
- Covert characters options:
 - No change – Do not change the case of the captured field value.
 - To lowercase - Convert to lowercase.
 - To uppercase case - Convert to uppercase.
- Allow only characters.
 - You specify the list of allowable character.
 - All others characters are eliminated.
- Filter strings.

- You specify the list of characters that must be removed.
- Replace.
 - Replace one text string with another.

Length & Value Content ruleset options

- Length.
 - Validate the number of characters or identified OMR boxes, such as check boxes, in the captured value of the current field.
 - Minimum of one character (**default** option for Form template).
- Value Content.
 - You provide an exact match value, a Partial Match value, or a Regular expression match value against which the recognized value is compared.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-78. Length & Value Content ruleset options

Length

- Validate the number of characters or identified OMR boxes, such as check boxes, in the captured value of the current field.
Minimum of one character (**default** option for Form template).
 - Minimum: Specify the minimum character length of the current field value.
 - Maximum: Specify the maximum character length of the current field value.
 - Minimum OMR: Specify the minimum number of identified OMR boxes to validate.
 - Maximum OMR: Specify the maximum number of identified OMR boxes to validate.

Value Content

- Validate that the value of the parameter that was entered matches the captured value of the current field. You provide comparison parameter options:
 - Exact match value.
 - Partial Match value.
 - Regular expression match value against which the recognized value is compared.

Date ruleset options

- Date.
 - Check date format.
 - Any valid date.
 - Range of dates.
 - Within # of days.
 - Reformat.

Examples: yyyy-mm-dd or mm/dd/yy

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-79. Date ruleset options

Date Ruleset Options

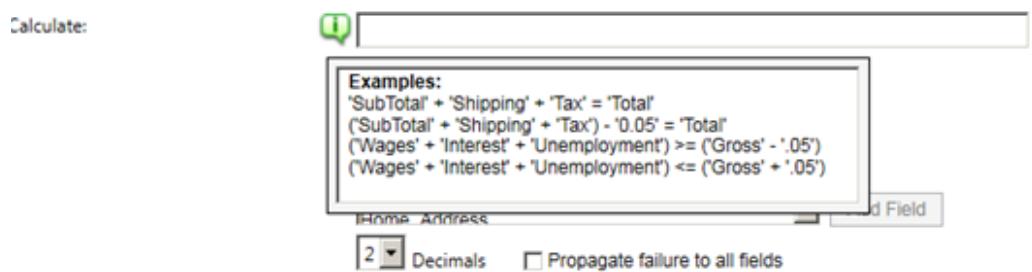
Validate that the captured value of the current field has an acceptable date format.

- Check date format.
 - Range of dates.
 - Check that the date is within the Range start and Range finish dates. (Use “Date start” and “Date finish” parameters)
 - Within # of days.
 - Check that the captured date lies within the “Number of days” value before or after the current date. (Use “Number of Days” property)
- Reformat.
 - You provide a new format that you want the date value to be converted to.



Currency & Numeric ruleset options

- Currency and numeric.
 - Sum name.
 - Percent numeric.
 - Currency format check box.
 - Less than or equal to.
 - Greater than or equal to.
 - Calculate.
 - Specify a formula for calculating the field value.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-80. Currency & Numeric ruleset options

Currency and numeric

Validate whether the captured value of the current field meets the currency format of the current locale.

- Sum name.

The name to use to locate the field or variable to sum the children of the current variable.

- Percent numeric.

Check that the characters in the captured value are n% numeric.

- Currency format check box.

Check whether the value meets the currency format specification for the locale.

- Less than or equal to.

The maximum value that is allowed for the field.

- Greater than or equal to.

The minimum value that is allowed for the field.

- Calculate.

Calculate the equation that is entered as a parameter and compares the result to the captured value of the current field object.

- Specify a formula for calculating the field value.

Lookup ruleset options

- **Lookup.**
 - **Database.**
 - Select the database to use for lookup.
 - **SQL statement.**
 - Look up the value to compare the field value against.
 - Populate fields from values returned from a database lookup.
 - **Populate field check box.**
 - This option specifies that you populate the field with the value from the database.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-81. Lookup ruleset options

Lookup

Validate the current field value by allowing it to be used in a database query, optionally populating fields with lookup.

- **Database:** Select the database to use for lookup
- **SQL statement**

Look up the value to compare the field value against.

Example: validate that a zip code really exists by looking it up in a database.

Populate fields from values returned from a database lookup.

Example: Populate the City and State fields based on the parameters returned from a zip code lookup operation.

- **Populate field check box**

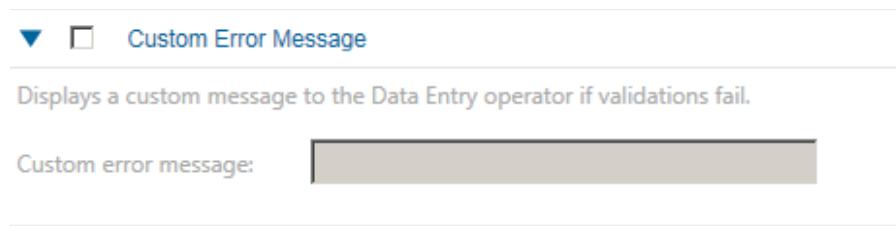
This option specifies that you populate the field with the value from the database.



Custom Error Message ruleset options

Display a custom message to the operator if the validation fails.

- Custom Error Message.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-82. *Custom Error Message ruleset options*

Exercise: Configure validate fields ruleset

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-83. Exercise: Configure validate fields ruleset

Exercise objectives

- Configure validate field rulesets
- Configure a new page fingerprint



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-84. Exercise objectives

Lesson 3.8. Export to IBM FileNet Content Manager



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-85. Export to IBM FileNet Content Manager

Lessons

- Introduction to FastDoc Datacap Server mode
 - Create an application with Form template
 - Create batch structure
 - Configure form template rulesets
 - Define fingerprints and test rulesets
 - Populate fields with keyword lookup
 - Configure validate field rulesets
-  Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-86. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system.
- This lesson describes how to configure the Form rulesets to export to an IBM FileNet Content Manager repository.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-87. Why is this lesson important to you?

Delivery options

- What does Export mean?
 - When Capture processing is complete, send documents and index parameters to a data repository or the next business process.
- Content Management Systems.
 - IBM Content Managers.
 - Third-Party Content Managers.
 - All CMIS compatible systems.
 - (See notes for details)
- Business-specific export options.
 - LOB applications.
 - Case and Business Process Management Workflows.
 - Databases and Electronic Data Interchange services.
 - (See notes for details)

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-88. Delivery options

IBM Content Management Systems

- IBM FileNet Content Manager

Use the FileNet P8 actions to export data to a FileNet Content Manager repository.
- IBM Content Manager

Use the IBMCM actions to upload documents into an IBM Content Manager repository.
- IBM Content Manager OnDemand (CMOD)

You can configure Datacap to export index data and files into Content Manager OnDemand.
- IBM FileNet IDM Export

Use the FileNetIDM actions to upload documents into an FileNet Image Services library.
- IBM Navigator on Cloud

Third-party Content Manager

- Microsoft SharePoint

Use the SPExport actions to upload documents to a Microsoft SharePoint library.
- Any CMIS repository

Content Management Interoperability Services (IBM CMIS) is an open standard.

Use the CMISClient actions to communicate with compliant content management systems over the internet.

Business-specific export options

- Line of Business (LOB) applications
- Case Based Applications
- Business Process Management solutions
- Electronic Data Interchange (EDI).

Other export options.

- Relational Databases

Use the ExportDB actions to set up and write information to an export database.

- XML Export

Datacap can export data to an XML file by using the actions in the ExportXML action library.

The ExportXML library includes actions that you can use to write data to an XML file.

- Local or network text file

Export configuration options

- FastDoc built-in Export Support:
 - Form and Learning template options:
 - A local text file with the Export ruleset. (Default)
 - IBM Content Manager with the “Export to IBM Content Manager” ruleset.
 - IBM FileNet Content Manager with the “Export to IBM FileNet Content Manager” ruleset.
- Export rulesets that are built with Datacap Studio from action libraries.
 - ExportDB action library exports to a connected database.
 - ExportXML action library exports as XML files.
 - CMISClient action library exports to CMIS compatible systems.
 - SPExport action library exports to Microsoft Share Point Manager.
 - FileNetIDM action library exports to FileNet Image Services Repository.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-89. Export configuration options

FastDoc built-in Export Support

There are four Export related Rulesets that are defined in Datacap that are used and are configurable with Ruleset UI from The FastDoc administrator interface.

- Export ruleset - Conventional ruleset that is configured to export text parameters to the application export folder.

Default Export configuration that is implemented in the Form and Learning templates is:

- Export to a local text file with the Export Ruleset.
- Use the Export actions to set up and write information to an export text file on the local system.
- The Export actions can set up line item values, batch, document, and page variables, path locations, before exporting the information.
- ExportAllFields action writes a comma-separated list of all index values to the text file.

Sample Text Output from the ExportAllFields action:

*,Mary Smith,245-67-8901,05-11-1966,Company A,,A12345678,2345 Maple Street,Smarter City,NY,10026
,Jose Flores,567-55-3456,06-05-1955,Company C,,C34567890,3434 Cypress Road,Smarter City,NY,10026*

- PreExport Compiled (UI) ruleset to convert output image to TIFF or PDF format.
- “Export to IBM Content Manager” Compiled (UI) ruleset to configure export to IBM Content Manager.
- “Export to IBM FileNet Content Manager” Compiled (UI) ruleset to configure export to IBM FileNet Content Manager.

Datacap Studio builds rulesets from action libraries.

- ExportDB action library exports to a connected database.
- ExportXML action library exports as XML files.

Use the ExportXML actions to set up and write information to an export XML file.

- CMISClient action library exports to CMIS compatible systems.
- SPExport action library exports to Microsoft Share Point Manager.
- FileNetIDM action library exports to FileNet Image Services Repository.

General export action

- In General each of the export action libraries mentioned in the previous slides, all have actions to do the following operations.
- Export actions:
 - Set up document attributes.
 - Build the record in memory.
 - Create folders and documents or data records.
 - Locate or open a connection to the file, database, repository, or business-specific process.
 - Authenticate with the destination server.
 - Upload documents to the server for storage or more processing.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-90. General export action

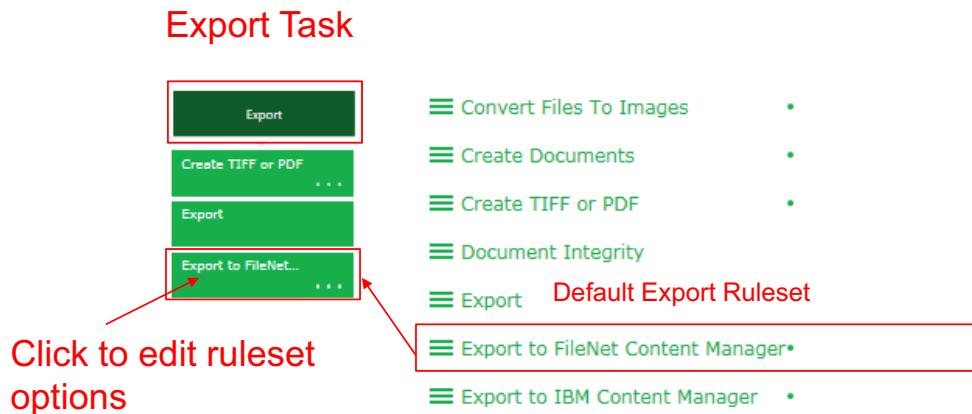
Help Path:

- Datacap > Datacap 9.0.1 > Reference > Action library summaries > Global actions
http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.dc.develop.doc/dcadt02.htm?lang=en



Select a Ruleset

- Form and Learning template options:
 - A local text file with the Export ruleset. (Default)
 - IBM FileNet Content Manager with the “Export to FileNet Content Manager” ruleset.
 - IBM Content Manager with the “Export to IBM Content Manager” ruleset.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-91. Select a Ruleset

The image on the slide shows the Export task and the Ruleset list from the Configure Workflow view of the FastDoc Admin interface.

- In the exercise for this Lesson, you configure the Export to FileNet Content Manager ruleset.
- Copy the Export to IBM Content Manager ruleset into the Export Task.
- Normally you would remove the Export ruleset, which writes index parameters to a text file but it is also OK to leave it in while you are debugging the application. Remove it before the application is released for production.
- The three dots in the lower right corner of the ruleset box, indicates that the ruleset is a compiled UI ruleset. This identification means that it has parameters that can be configured with FastDoc



Configure the Batch parameters

- Export to FileNet Content Manager ruleset has three levels of parameters to configure:
 - Batch level
 - Document level
 - Field level

Batch Information

FileNet Content Manager URL:*	<input type="text" value="http://ecmedu01:9080/wsi/FNCEWS40MTOM/"/>
User ID:*	<input type="text" value="p8admin"/>
Password:*	<input type="password" value="*****"/>
Confirm password:*	<input type="password" value="*****"/>
Locale:	<input type="text" value="English (United States)"/>
Storage object id:*	<input type="text" value="DCExport"/>
Parent folder:	<input type="text" value="/Beneficiary Forms/"/>
Subfolder to create for batch:	<input type="text" value=""/>
Number of upload attempts:	<input type="text" value="0"/>
Upload timeout:	<input type="text" value="600000"/>

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-92. Configure the Batch parameters

The image on the slide shows one image that indicates that the Export to FileNet Content Manager ruleset has three levels of configuration option. The second image shows the configurable options for the batch level parameters.

- The following parameters are Custom Values and must be configured in the Datacap Application Manager.
 - FileNet Content Manager URL - FileNetURL
 - User ID* - FileNetUserID
 - Password – FileNet Password
- Storage object ID: defines the Repository name.
- Parent Folder: defines the folder name within the repository.
- Subfolder to create for batch: defines a subfolder that can be provided for each batch.



Configure the Document parameters

- The upload file name must match the document ID.
- Document class ID must match the repository DocClass.
- Document file extension tif is assumed unless overridden here.
- See notes comment about Add Rows

FastBuild

Document

Bene_Document

Bene_Page

Other

Document Information

The Bene_Document document will be exported.

Document title:*	@ID
Document class ID:*	Beneform
Document file extension:	pdf

Document properties:

Symbolic name	Value	Type	Multi

Add Row

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-93. Configure the Document parameters

The image on the slide shows the configurable options for the document level parameters.

Document title:

- The FNP8_Upload action requires that the upload file names match the document ID.
- The file name must match @ID.pdf or @ID.tif, where @ID is the ID of the document level.
- The document file for export must exist before calling the FileNet export actions.

Document class ID:

- All index parameters must be matched to the properties defined in the IBM FileNet Content Manager document class ID defined here.
- If Index parameter names do not match Document class property names then, they must be mapped to the property symbolic name.

Document file extension:

- This field is used to append to the document name to match the name of the document that must be uploaded.
- If the field is blank, then TIF is assumed

Document properties: Add Row control

- The document level option is useful for exporting document variables, which are document level metadata.
- Document variable values come from various sources:
 - Data that is entered by an operator when the entire batch was scanned.
 - Data that persists from document to document and many others.
- The syntax for referring to page level parameters is as follows.

`@D\myPageType\myFieldType`

Note: It is not recommended that you define page level parameters for export at the document level by using the Add Row control. The preferred method is to define the parameter for export at the field level. See the next slide.



Configure the Field parameters

- This is the preferred method of exporting page parameters.
 - Click the Field Information check box to enable export for the field.
 - The Symbolic name must match the Document Class Symbolic Name.
 - The property type must match the Document Class property type.
 - Multi-value property is used to identify array type properties.



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-94. Configure the Field parameters

The image on the slide shows the configurable options for the Field level parameters.



View IBM Repositories with Datacap Navigator

- IBM Content Navigator Supported Repositories

The screenshot shows the 'IBM Content Navigator' interface. On the left, a sidebar lists various options: Desktops, **Repositories** (which is selected and highlighted with a red box), Sync Services, FileNet Content Manager, Content Manager OnDemand, Daeja ViewONE, Viewer Maps, and Plug-ins. The main area has tabs for 'Desktops' and 'Repositories'. Under 'Repositories', there is a toolbar with 'New Repository' (selected and highlighted with a red box), Edit, Delete, and Refresh buttons. A dropdown menu under 'New Repository' lists several repository types: Content Manager, Content Manager OnDemand, FileNet Content Manager, Content Management Interoperability Services (CMIS), and Datacap Application. The 'FileNet Content Manager' option is highlighted with a red box.

- `http://<navigator server name>:9080/navigator > Repositories`
- Configure Repository access.
 - For visibility into the supported repositories.
- Configure Datacap application access.
 - For Navigator process, and monitor control of Datacap applications.

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-95. View IBM Repositories with Datacap Navigator

The image on the slide shows the supported repository types in the New Repository selection list.

To configure access to a repository from IBM Content Navigator (ICN):

1. Open a Browser window and browse to `http://<navigator server name>:9080/navigator`.
`<navigator server name>` = ecmedu01 for the class image.
2. Log in as navigator administrator. (p8admin/IBMFContentManagerP8) for class image.
3. Click the Repositories option in the left side toolbar.
4. Click New Repository.

Configure Repository access

The supported repositories are:

- Content Manager – IBM Content Manager (CM8)
- Content Manager On Demand – CMOD
- FileNet Content Manager – IBM FileNet Content Manager (P8)
- Content Manager Interoperability Services (CMIS)

Configure Datacap application access

- Datacap Application

- Each application that Datacap Navigator controls must be defined by using the Datacap Application option on the repositories view.
- You must connect to the repository before you configure parameters, system Properties, Folders, and so on.



IBM

IBM FileNet repository configuration parameters

- IBM FileNet Content Manager repository configuration parameters.

FileNet Content Manager repository: **DCEExport**

General		Configuration Parameters	System Properties	Browse	Search	...
You must connect to the repository before you configure parameters, system properties, folders, and so on.						
* Display name: ?	DCEExport					...
* ID: ?	DCEExport					...
* Protocol:	EJB					...
* Server URL: ?	iiop://ecemedu01:2809/FileNet/Engine					...
* Object store symbolic name: ?	DCEExport					...
* Object store display name: ?	DCEExport					...
Connect...						...

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-96. IBM FileNet repository configuration parameters

The image on the slide shows the required parameters for configuring a connection to an IBM FileNet Content Manager repository.

Parameters are:

- Display name: DCEExport
- Server URL: iiop://ecemedu01:2809/FileNet/Engine
- Object store symbolic name: DC Export
- Object store display name: DC Export



View Exported objects in the repository

- Use the Datacap Navigator desktop to view export objects in the IBM FileNet Content Manager repository.

Select repository view
Repository name

Folder name

Document object

Document image

Document Index properties

Properties

Name	Size	Modified By	Modified On	Major Version	Employee_Name
ECMEDU01.20150325.000008.01	60 KB	P8Admin	3/25/2015, 9:13 PM	1	Mary Smith
ECMEDU01.20150325.000008.02	60 KB	P8Admin	3/25/2015, 9:13 PM	1	Jose Flores
ECMEDU01.20150325.000008.03	61 KB	P8Admin	3/25/2015, 9:13 PM	1	Ahmed Rizwan

© Copyright IBM Corporation 2017

Figure 3-97. View Exported objects in the repository

The image on the slide shows the view of an IBM FileNet Content Manager repository.

You use the Datacap Navigator desktop to view the repositories.

1. Open a Browser window and browse to `http://<navigator server name>:9080/navigator/?datacap&metaie=10`.
`<navigator server name>` = ecmedu01 for the class image.
2. Log in as navigator administrator. (admin/admin) for class image.
3. Click the Repositories option in the left side toolbar.
4. Double-click the folder used Datacap export. For this lesson Beneficiary Forms

Exercise: Export to IBM FileNet Content Manager

Requires:
Course Exercise Guide
Student system

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-98. Exercise: Export to IBM FileNet Content Manager

Exercise objectives

- Export to IBM FileNet Content Manager



FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-99. Exercise objectives

Unit summary

- Create an application with Form template
- Create batch structure
- Configure form template rulesets
- Define fingerprints and test rulesets
- Populate fields with keyword lookup
- Configure validate field rulesets
- Export to IBM FileNet Content Manager

FastDoc Datacap Server mode

© Copyright IBM Corporation 2017

Figure 3-100. Unit summary

Unit 4. Datacap Studio Introduction

Estimated time

01:00

Overview

This unit introduces you to the capabilities Datacap Studio Application Building capabilities.

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/datacap_9.0.1.htm

Unit Objectives

- After completing this unit, you should be able to:
- Describe the components in Datacap Studio for application development process
- Use the Datacap Studio

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-1. Unit Objectives

Lesson 4.1. Datacap Studio – Rulemanager



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-2. Datacap Studio – Rulemanager

Lessons

-  Datacap Studio – Rulemanager
 - Datacap Studio – Zones

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-3. Lessons

Why is this lesson important to you?

- As an Application Developer, you build and deploy applications with the Datacap capture system.
- To do these tasks effectively, you must be familiar with the Datacap Studio interface.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-4. Why is this lesson important to you?

Panel Organization within Datacap Studio

- Three Main Tabs
 - Rulemanager, Zones, and Test Tabs.
- Rulemanager Tab
 - Contains 5 panels where you view document structures, rulesets, rules, functions, and task profiles.
 - Can also be used to extend application created in Datacap FastDoc.
- Zones Tab
 - Contains 4 panels where you add and define fingerprints and view and modify properties of selected objects.
- Test Tab
 - Contains 8 panels in which you can view information and properties of batches, jobs, documents and rulesets.
 - Is a test environment that allows you to debug your application, set breakpoints, step through and skip pieces of your application, and lots more functionality.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-5. Panel Organization within Datacap Studio

Panel Organization within Datacap Studio

Datacap Studio has three views that you select with the three tabs across the top of the window.

- Rulemanager
- Zones
- Test

The Rulemanager View

This pane is the primary application development area. The default layout of the Rulemanager view for Datacap Studio window has four panes. One of these panes has two functions selectable by tabs at the top of the pane. You can customize the pane layout by removing or adding panes or by moving them to different locations. The Rulemanager panes are:

- Document hierarchy
- Rulesets
- Action library
- Task profiles
- Properties

The Zones View

Contains 4 panels where you add and define fingerprints and view and modify properties of selected objects.

The Test Tab

Contains 8 panels where you can view information and properties of batches, jobs, documents and rulesets.

It also functions as a test environment that allows you to debug your application by setting breakpoints,

stepping through and skipping blocks of your application, along with many more debugging capabilities.

Document hierarchy

The document hierarchy describes the structure of the documents that your application is designed to handle.

The levels within the hierarchy are:

- Batch
- Document
- Page
- Field.

Rulesets

- A ruleset consists of one or more rules.
- Rulesets are assigned to tasks in the task profiles.
- Rules are assigned to process specific objects in the document hierarchy (for example, to analyze each page and identify its type).

There are two Types of Ruleset:

- Conventional rulesets
- Compiled or UI rulesets

Task profiles

A workflow consists of a series of tasks and defines a way to process documents. Datacap applications can include multiple workflows. Each task is linked to a task profile and the task profile determines the order that tasks are processed.

Actions library

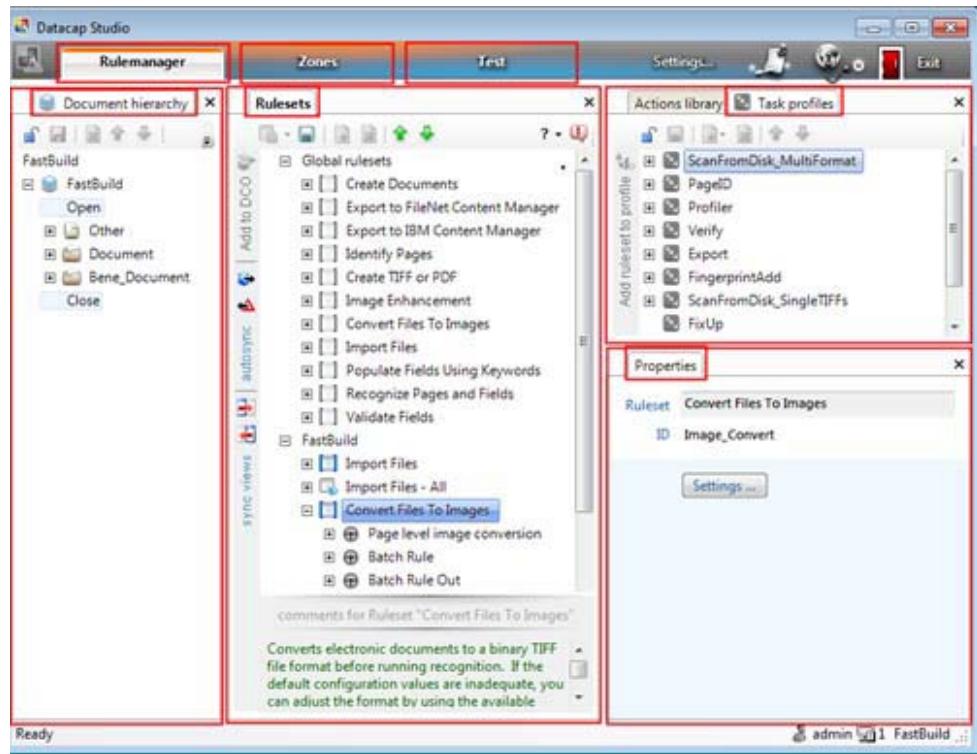
Provides access to the complete library of prebuilt actions. An action is the smallest component of a workflow. It is the equivalent of an instruction in programming terms. To get help on an action, select the action and click Information.

Properties

Shows the properties for the selected document hierarchy or ruleset object. If the corresponding pane is locked for editing, you can also modify existing properties, including specifying action parameters.



Datacap Studio Overview



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-6. Datacap Studio Overview

Datacap Studio overview

- Datacap Studio is a full function rich application development environment for Datacap Capture.
- It also includes an Application Wizard that gives you a head start on application development by generating a basic application framework, complete with the supporting folders.
- It provides the tools that you need to develop and test your application.
- Any application that is configured in the FastDoc rapid development environment can be developed further in Datacap Studio to achieve functions that are not supported in FastDoc.
- Up to Datacap 8.1 it was the only application development tool.
- In Datacap 8.1, FastDoc was a limited functioning rapid application development tool.
- In Datacap 9.0, FastDoc abilities are vastly expanded.

Start Datacap Studio

To open a sample application in Datacap Studio:

1. Click Start > All Programs > IBM Datacap Developer Tools > Datacap Studio.
2. In the Select Application window, select one of the existing sample applications (for example, FastForm) and click Next.
3. In the Taskmaster Login window, make sure that the NT authentication check box is not selected.
4. Enter User ID, Password, and Station ID and then click Finish. (Default Credentials: User ID = admin, Password = admin, Station ID = 1)

Rulemanager View

- Document hierarchy
 - Describes the structure of the documents your application is designed to handle.
- Rulesets
 - Rulesets are assigned to tasks in the Task profiles.
 - Rules are assigned to process specific objects in the document hierarchy.
- Task profiles
 - Lists the workflow of tasks. Determine the order that rules are processed while processing a task.
- Actions library
 - Provides access to the complete library of prebuilt actions.
- Properties
 - Shows the properties for the selected document hierarchy or ruleset object.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-7. Rulemanager View

Document hierarchy

The document hierarchy describes the structure of the documents that your application is designed to handle.

The levels within the hierarchy are **Batch**, **Document**, **Page**, and **Field**.

Ruleset

A ruleset consists of one or more rules and is assigned to tasks in the task profiles.

Rules are assigned to process specific objects in the document hierarchy (for example, to analyze each page and identify its type).

There are two types of rulesets:

- Conventional rulesets
- Compiled or UI rulesets

Task profiles

A workflow consists of a series of tasks and defines a way to process documents. Datacap applications can include multiple workflows. Each task is linked to a task profile and the task profile determines the order that tasks are processed.

Actions library

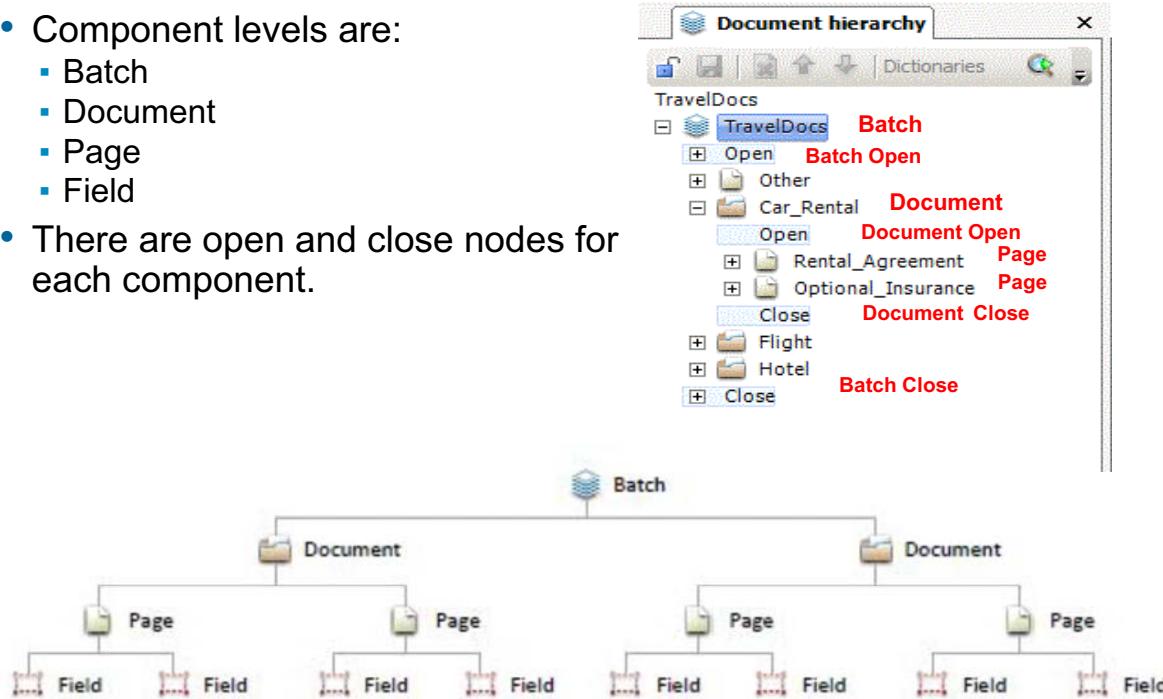
A complete library of pre-built actions. An action is the smallest component of a workflow. It is the equivalent of an instruction in programming terms. To get help on an action, select the action and click Information.

Properties

Shows the properties for the selected document hierarchy or ruleset object. If the corresponding pane is locked for editing, you can also modify existing properties, including specifying action parameters.

Document Hierarchy (DCO)

- Component levels are:
 - Batch
 - Document
 - Page
 - Field
- There are open and close nodes for each component.



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-8. Document Hierarchy (DCO)

The images on the slide illustrate:

- Document Hierarchy view.
- A graphical representation of the Batch Structure Batch, Document, Page, Field.

Document Hierarchy

The Document hierarchy is also known as the DCO. In FastDoc the DCO is also known as the Batch Structure.

The component levels within the hierarchy are:

- **Batch**

At the top of the document hierarchy, is the batch, which refers to all pages of all document types.

- **Document**

Beneath the batch level, the document hierarchy defines:

- The document types that your application can process. You might have only one type, or you might have multiple types.

Example: TravelDocs application processes car rental documents, airline ticket, humanitarian relief fund expense forms, and donation receipts.

- **Page**

- Other page type

There is always a page type of Other associated with the batch. This page type is used to temporarily assign a page type to all incoming pages until each page is identified as being one of the recognized pages in the application.

- Custom pages that are associated with a defined document.

The page types define the pages within each document type. Each document might have only one page type, or it might have multiple types. The page definitions define the number and order of pages within each document type. Pages can be required or optional.

Example: The car rental document includes the rental agreement page and the optional insurance page, while the flight document has only an air ticket page.

Example: A car rental document has a maximum of two pages. The rental agreement page is required and must come first. The insurance coverage page is optional.

- **Field**

- The field nodes of the DCO are the data fields within each page type. Data fields can be required or optional.
 - Data fields configured for extracting data from the images are mapped to the fields defined in the DCO.

- **Open and Close nodes.**

Each component has an Open and a Close node in the hierarchy. The Open and Close nodes are used to assign processing rules to the beginning or the end of each component.

Rules and Rulesets

- **Rule**
 - Entity that is tied to a DCO object.
 - An ordered set of functions that processes an object until a function is evaluated as True or all the other functions are exhausted.
 - Reusable and extensible.
- **Ruleset**
 - A set of rules that is used to do a particular type of processing on various DCO objects, i.e. the Validate Ruleset.
- **Task Profile**
 - A task profile is made up of a number of rulesets that are arranged in a particular sequence to produce the desired task processing results.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-9. Rules and Rulesets

Rule

A rule is the definition of some processing to be done and is tied to a DCO object. It is made up of an ordered set of functions that processes an object until a function is evaluated as True or all the other functions are exhausted. It is reusable and extensible.

Rulesets

A ruleset consists of one or more rules and is assigned to tasks in the task profiles. Rules are assigned to process specific objects in the document hierarchy (for example, to analyze each page and identify its type). There are two types of rulesets in Datacap 9.0:

- Conventional rulesets
- Compiled or UI rulesets

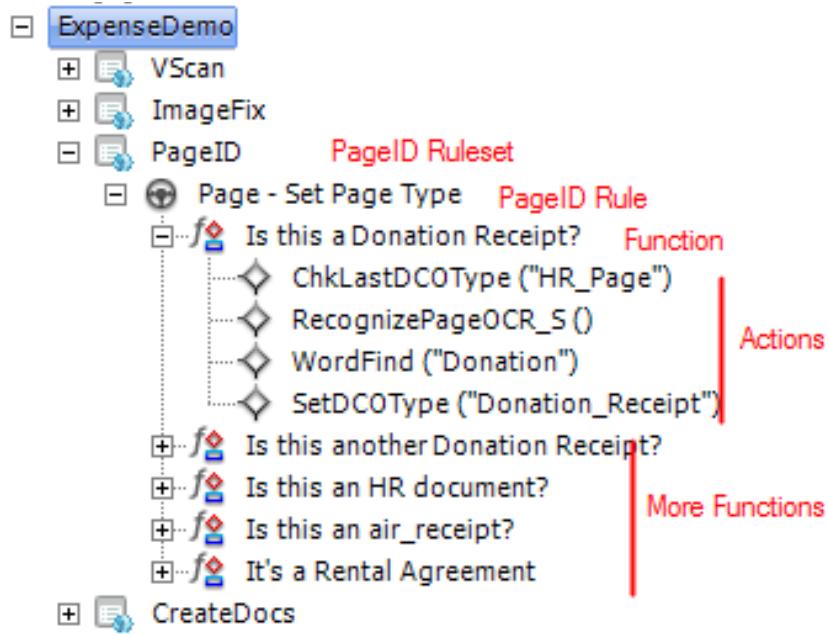
Task profiles

A workflow consists of a series of tasks and defines a way to process documents. Datacap applications can include multiple workflows. Each task is linked to a task profile and the task profile determines the order that tasks are processed.



Rules and Rulesets

- Rulesets and Rules:



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-10. Rules and Rulesets

The image on the slide illustrates a sample ruleset, showing the structure of Batch, Rulesets, Rules, Functions, Actions.

Rulesets

A ruleset consists of one or more rules. In the following example, the default PageID ruleset has one rule: Page – Set page Type. You can see the rules that are associated with each ruleset in the Rulesets pane on the Datacap Studio - Rulesets view.

Rules

Rules are assigned to process specific objects in the document hierarchy (for example, to analyze each page and identify its type). However, you define the rule by the programmed functions and actions that you put in it. The default PageID rule consists of five functions and multiple actions in each function, as shown in the screen capture.

The PageID function first runs the ChkLastDCOType("HR_Page") action. If ChkLastDCOType is successful (returns True), the function runs RecognizePage(OCR_S()).

If ChkLastDCOType fails (returns False), the function fails and Datacap Capture runs the next function within the rule.

If processing reaches the end of a rule and no error is encountered then the true condition is returned and Datacap continues to process the next rule for the current object.

Add Rules to the DCO.

See Rule mapping in subsequent slides.

Actions and Functions

- Actions
 - Fundamental building block from which rules are built.
 - Built with .Net functions or VB script functions.
 - Always return True or False to the Rulerunner service.
 - A group of actions forms a function.
- Functions
 - Ordered list of actions that run until one of the actions returns the value of False.
 - A rule can have multiple functions that will be executed in the listed order until one of them completes, or all functions have failed.
 - If all functions fail, the rule did not execute successfully.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-11. Actions and Functions

Actions

Actions are the fundamental building blocks that are used to define rules. They always return True or False to the Rulerunner service and may perform other types of actions. A group of actions forms a function.

Functions

A function is an ordered list of actions that will run until one of the actions returns a value of False. Rules are made up of functions that will execute in the listed order until one of the functions completes or all functions have failed.

Actions library

Provides access to the complete library of prebuilt actions for a particular type of functionality. An action is the smallest component of a workflow. It is the equivalent of an instruction in programming terms. To get help on an action, select the action and click Information.

Properties

Shows the properties for the selected document hierarchy or ruleset object. If the corresponding pane is locked for editing, you can also modify existing properties, including specifying action parameters.



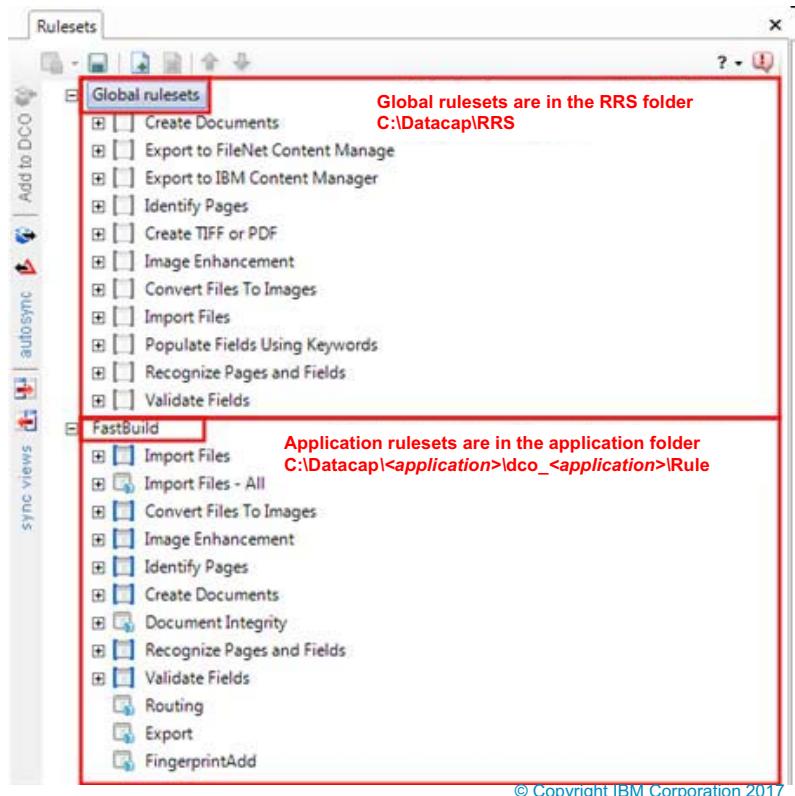
Global and Application Rulesets

- Global rulesets:

- RRS folder

- Application rulesets:

- Application folder



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-12. Global and Application Rulesets

The image on the slide illustrates the view of the Datacap Studio Ruleset pane :

- It shows the Global rulesets in the upper section
- It shows the Application rulesets in the lower section

Global rulesets:

- Are a pool of built-in rulesets included with the Datacap release.
- Are in the C:\Datacap\RRS folder.
- Are typically but need not necessarily be compiled (UI) rulesets.
- They are ruleset templates and cannot be configured while in the Global ruleset group. Icon is gray (Not configurable)
- A copy must be placed into the application before the application copy can be configured.
- Developers can copy customized application rulesets into the RRS folder. This makes them global and available for other developers to share.

Application rulesets:

- Are the rulesets that are selected to build the application.
- Are in the C:\Datacap\<application>\dco_<application>\rules folder.
- Might be a copy of a global ruleset.
- Can be conventional or compiled rulesets.
- Icons are not gray but have color. (Configurable)
- For compiled (UI) rulesets, you can access and configure the properties that are configurable from through the FastDoc UI interface.
- Compiled rulesets cannot be locked and structurally modified through Datacap Studio

Note: The Global rulesets that are in the RRS folder cannot be configured because they are templates. They must be copied into the application folder by doing a copy and paste before they become configurable.

Ruleset Types

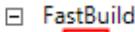
- Compiled UI rulesets:
 - Have a user interface to make parameter configuration easier.
 -  FastBuild
 - +  Validate Fields
 - Have a symbol that is bracketed by vertical bars.
- Conventional rulesets:
 - Have no user interface to make parameter configuration easier.
 - +  Document Integrity
 - Have the traditional ruleset icon.

Figure 4-13. Ruleset Types

The images on the slide illustrate:

- The icon used to represent compiled UI rulesets
- The icon used to represent conventional rulesets

Compiled rulesets:

- Are self-contained application building blocks.
- Reduce the expertise that is needed to create applications.
- Are used in the built-in Forms and Learning templates.
- Make applications more consistent and easier to maintain.
 - Done by standardizing how core functions are implemented.
- Can be easily included into an application by an Application Builder.
- Have two files that represent the ruleset.
 - ValidateFields.Rul.dll - The .Rul.dll files contain the rule, function, and action definitions.
 - ValidateFields.Rul.dll.config - The .Rul.dll.config files are XML files that contain the settings for the user interface.

- Configuration can be done through Settings either with FastDoc or Datacap Studio.

Conventional rulesets:

- Rulesets that are created in legacy applications that are created before Datacap 9.0.
- Custom rulesets that are created with Datacap Studio to do specific functions that cannot be accomplished with the built-in Compiled (UI) rulesets.
- Have access to hundreds of actions that can be used to create powerful capture solutions.
- Can also be used in the built-in Forms and Learning templates.
- Can be easily included into an application with FastDoc when it is created and configured in Datacap Studio.
- Have one file that represents the ruleset.
 - ValidateFields.Rul.dll - The .Rul.dll files contain the rule, function, and action definitions.
- Configuration can be done only by setting the action parameters with Datacap Studio.

Rule Mapping

- Rule mapping can be done only by using Datacap Studio.
- There are two methods of doing rule mapping:
 - Legacy mapping method:
 - Rule mapping data in legacy applications is stored in the dco_<application>.xml file.
 - Datacap 9.0 mapping method:
 - Rule mapping data in Datacap 9.0 applications is stored in the <ruleset>.Rul.dll.config file.
 - This method makes it easier to share rulesets with other applications or between developers.
 - This method greatly simplifies application management.
 - Manual configuration or need to develop rules to do application management now happens automatically.
 - Mappings reduce the need for learning how to configure routing and other more technical issues as Datacap 9.0 is enhanced.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-14. Rule Mapping

Legacy rule mapping method:

To map a rule to a DCO object:

1. Select the object in the Document hierarchy pane that you must add the rule to.
2. Then, select the rule to add.
3. Next, click the Add to DCO link on the left of the Ruleset pane.
4. The default behavior is to add the rule to the Open node of the selected DCO object.
5. To add a rule to the Close node, you must explicitly select the Close node in the Document hierarchy pane before clicking Add to DCO.

Datacap 9.0 mapping method:

To map a rule to a DCO object:

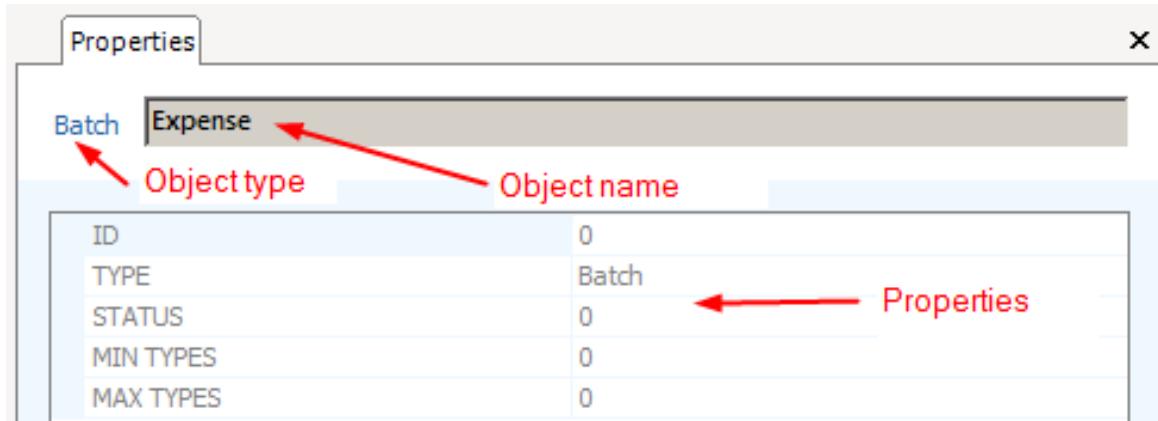
1. Expand a compiled (UI) ruleset
2. Select the rule that you want to map.
3. The mapping rule options are shown in the Properties pane.
4. Expand either “Run rule at start of ...” or “and at end of ...” view.

5. Select the Object type that you want to map the rule to.
6. If you want to map to a specific object of the selected type, click the + symbol and select the specific object.



Properties tab

- View object properties
- In general the properties pane shows the following:
 - Object type
 - Object name
 - Object properties, static and configurable. (If applicable)



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-15. Properties tab

The image on the slide illustrates the view of the Properties tab on the Datacap Studio Rulemanager window.

Properties

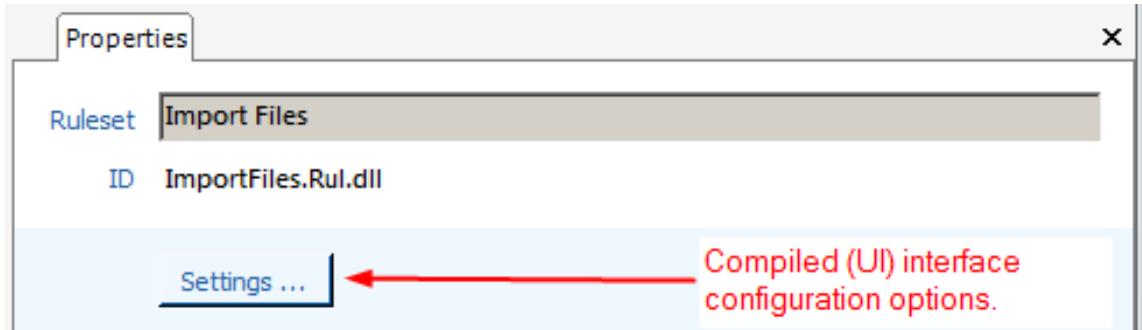
- Shows the properties for the selected document hierarchy or ruleset object. If the corresponding pane is locked for editing, you can also modify existing properties, including specifying action parameters.
- Each object type in the Document hierarchy and the Rulesets pane have a unique representation in the Properties pane.
- Document hierarchy object types are:
 - Application
 - Batch
 - Document
 - Page
 - Field

- Ruleset object type properties are:
 - Application
 - Ruleset
 - Rule
 - Function
 - Action
- On the following slides you look closer at the properties view for Ruleset, Rule, Function and Action objects.



Ruleset object properties tab

- Configure UI Settings
 - For compiled (UI) rulesets, the Settings ... control opens the UI interface for setting the Ruleset configuration options.
 - The Settings UI view is the same view seen in FastDoc for setting Ruleset configuration options.



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-16. Ruleset object properties tab

The image on the slide illustrates Properties view of the Datacap Studio window when a ruleset is Selected in the ruleset pane. It shows the settings control that opens the Compiled UI interface configuration options.



Rule object properties tab

- Map Rule to DCO objects

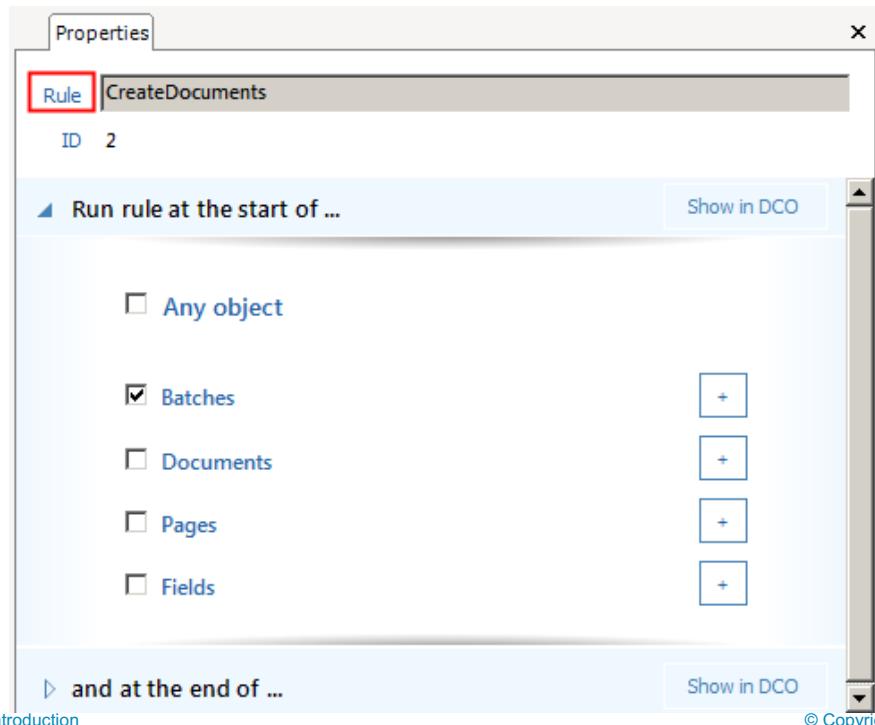


Figure 4-17. Rule object properties tab

The image on the slide illustrates Properties pane of the Datacap Studio window when a rule is selected in the Ruleset pane. It shows the controls used to associate rules with the DCO or Batch Structure objects.

Conventional Rule Mapping Method

- Conventional Ruleset rules are mapped using the pre 9.0 method.
 - Select the Rule in the Ruleset pane, select the DCO object in the Document hierarchy pane, and click the Add to DCO control.
 - Using the conventional method the mappings are visible in the Open or Close node for DCO object.

Datacap 9.0 rule Mapping Method

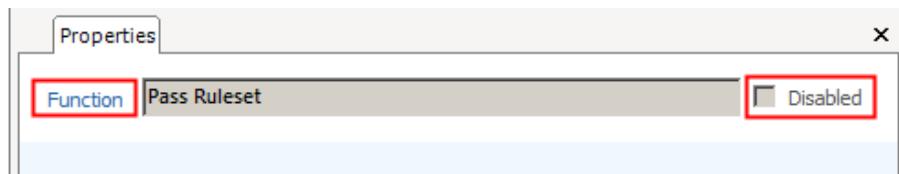
- Only for Compiled (UI) ruleset rules.
 - Rule mapping can only be done this way (as shown in the slide graphic) for Compiled (UI) Rulesets.
- Mapping at beginning or end of object rule processing.
 - Mapping configuration done under 'Run rule at the start of ...' is equivalent to mapping a rule to the Open node in the conventional method.

- Mapping configuration done under 'and at the end of ...' is equivalent to mapping a rule to the Close node in the conventional method.
- Mapping to all objects or all objects a a specific type.
 - The 'Any object' check box causes the selected rule to associate with all DCO objects.
 - The 'Batch' check box causes the selected rule to associate with all batch level objects.
 - The 'Documents' check box causes the selected rule to associate with all document level objects.
 - The 'Page' check box causes the selected rule to associate with all page level objects.
 - The 'Field' check box causes the selected rule to associate with all Field level objects.
- Mapping to a specific object of a specific type.
 - To map to a specific object instead of all objects of one type, click the + symbol and select the DCO object from the object list that you are presented with.

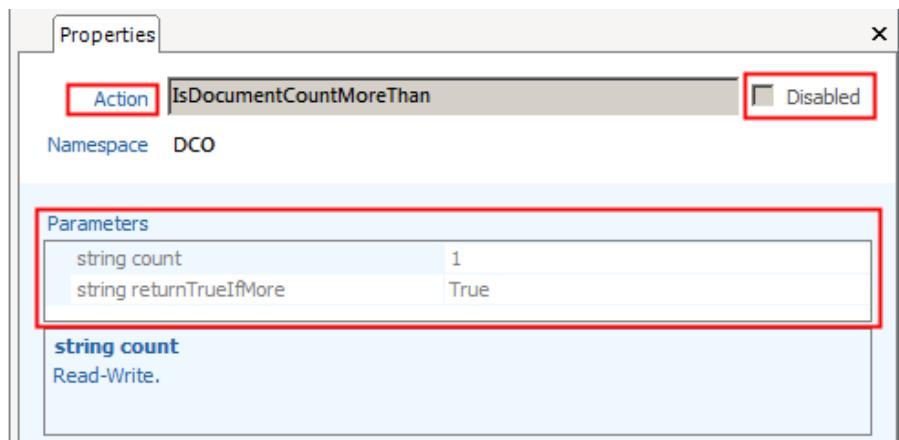


Function and Action object properties tab

- Disable the Function



- Disable the Action and configure action parameters



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-18. Function and Action object properties tab

The images on the slide illustrate the properties pane for functions and for actions that allow for disabling the selected object.

Disable Functions and Actions.

- The disable Function and Disable action options are to enable the activation or deactivation of UI Ruleset options.
- It can also be used for debugging purposes.

Set Action Parameters.

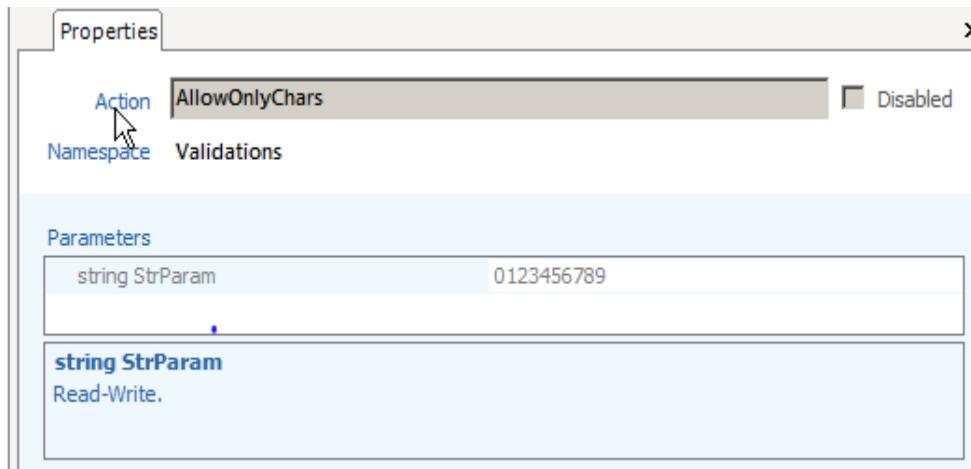
- The parameters that are required by the Actions are provided in the parameters area of the Properties pane.



IBM

Action object properties tab

- View object properties



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-19. Action object properties tab

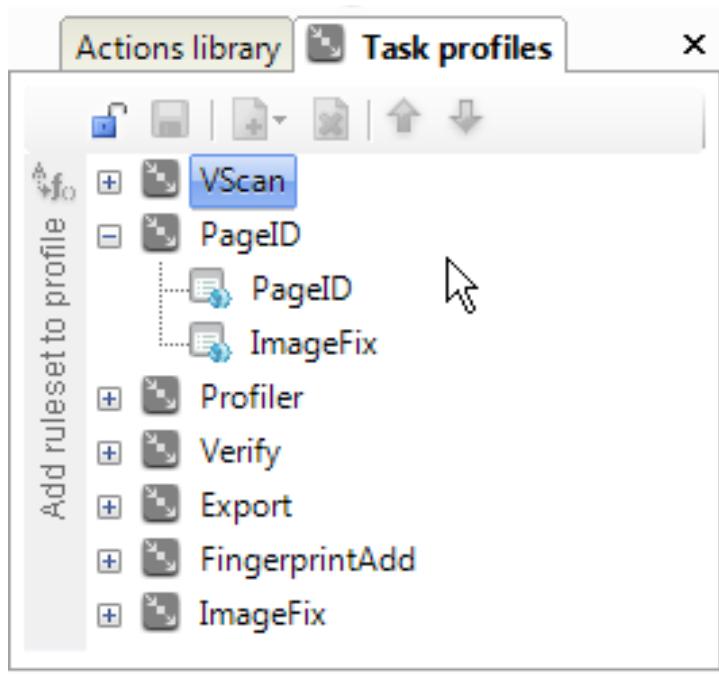
Properties

Shows the properties for the selected document hierarchy or ruleset object. If the corresponding pane is locked for editing, you can also modify existing properties, including specifying action parameters. The properties pane is where you go to set the parameters that are passed to actions.



Task Profiles

- Tasks and Rulesets



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-20. Task Profiles

The image on the slide illustrates the task Profiles pane on the Datacap Studio Rulemanager window.

Task profiles

A workflow consists of a series of tasks that defines a way to process documents. Datacap applications can include multiple workflows. This class focuses on single workflow applications.

A standard workflow for processing documents from Datacap Desktop takes a batch of documents through each of the processing steps.

A job consists of one or more tasks. To process a batch of documents, you must run the batch through each task in the selected job. Some tasks (for example, Export) run without operator intervention, but others (for example, Verify) require an operator.

The job type that you select determines the tasks in the workflow. You can see the tasks that are associated with each job type by looking in the Workflow pane on the Datacap Studio Zones view. A typical workflow for a job includes five tasks:

- VScan
- PageID
- Profiler
- Verify
- Export

Each task is linked to a task profile and the task profile determines the order that tasks are processed.

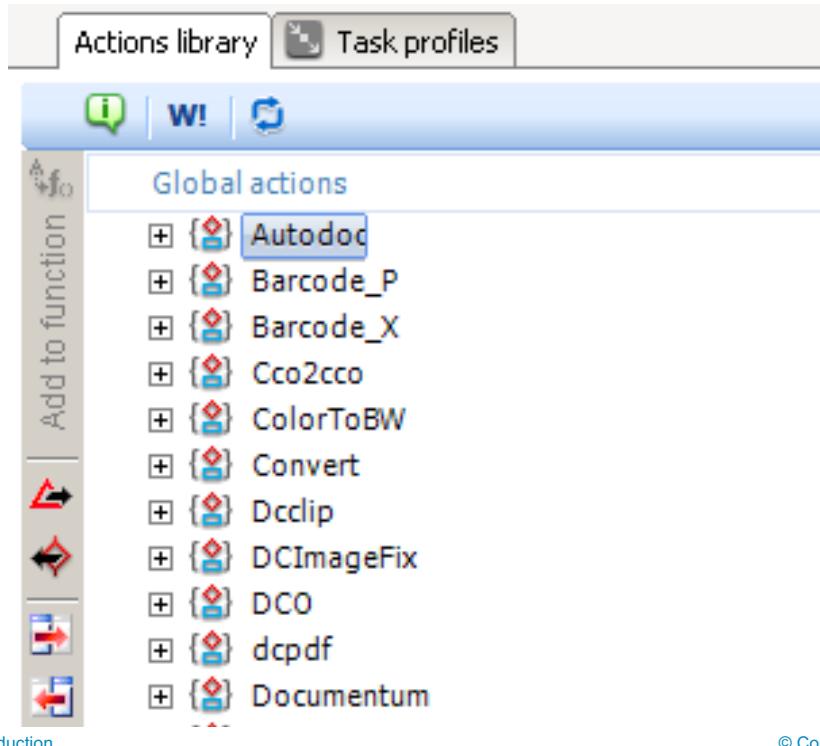
Add Rulesets to the Task profile.

Any new task that is defined must be added to a Task in the Task profiles for it to be included in the workflow for a batch. To add a ruleset to a task in the Task profile, you must select the ruleset in the Rulesets pane. Next, you must select the task in the Task profile pane that you want to be added to the selected task profile. Then, click the 'Add ruleset to profile' link on the left of the Task profiles pane.



Actions Library

- Tasks and Rulesets



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-21. Actions Library

The image on the slide illustrates the Actions library pane of the Datacap Studio Rulemanager window.

Actions library

Provides access to the complete library of prebuilt actions. To get help on an action, select the action and click Information. Each item that is listed in the Action library pane is a group of actions of a particular type. Click the plus sign next to each group to expand and show all the actions in the group.

Add actions to rules.

When building rules in the Ruleset pane, actions from one or more Action Libraries can be added to the functions. One or more functions makes a rule. To add actions to a rule in the Ruleset pane, select the function in the Ruleset pane that must receive the action. Next, you select the action to add to the rule. Then, click the Add to function link on the left of the Action library pane.

Exercise: Explore Datacap Studio - Rulemanager

Requires:
Course Exercise Guide
Student system

Figure 4-22. Exercise: Explore Datacap Studio - Rulemanager



Exercise objectives

- Explore the Datacap Studio - Rulemanager

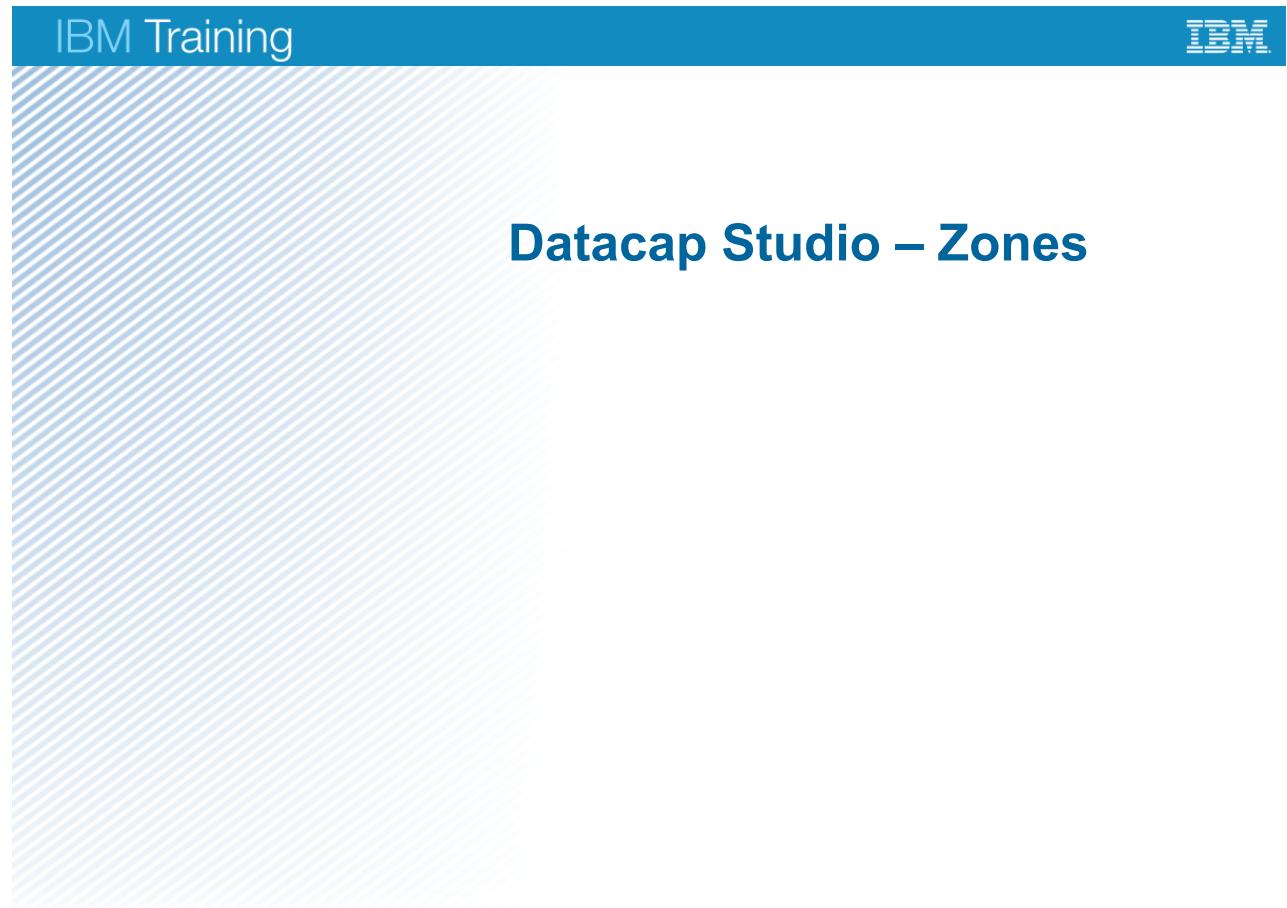


Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-23. Exercise objectives

Lesson 4.2. Datacap Studio – Zones



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-24. Datacap Studio – Zones

Lessons

- Datacap Studio – Rulemanager
- ▶ Datacap Studio – Zones

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-25. Lessons

Why is this lesson important to you?

- As a Application Developer, you build and deploy applications with the Datacap capture system.
- To do these tasks effectively, you must be familiar with the Datacap Studio interface.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-26. Why is this lesson important to you?

Zones View

- Fingerprints – Displays the application's fingerprint library.
- Document hierarchy – Defines the structure of the documents you are processing and how each element within the structure is processed.
- Properties – You can manage the properties for the selected document hierarchy object. It also lets you specify recognition options for the selected object.
- Image View – Shows the selected fingerprint image and any recognition zones.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-27. Zones View

The Zones View is where you create page fingerprints and set up recognition zones. The Zones view includes the following panes:

Fingerprints

Shows the application fingerprint library and lets you add fingerprints for new page types.

Document hierarchy

Defines the structure of the documents that you are processing and how each element within the structure is processed.

Properties

Shows the properties for the selected document hierarchy object. If the document hierarchy is locked for editing, you can also modify existing properties.

The Properties pane also lets you specify recognition options for the selected object. Datacap Capture supports multiple recognition engines. Panes for ICR/C, BAR/P, and OCR/S are shown by default. You can access other panes by right-clicking within the Properties panel and selecting Show panes.

Image view

Shows the selected fingerprint image and any recognition zones. This view is where you draw new recognition zones. If you created the fingerprints with full page recognition, the Text pane at the bottom lets you view the recognition results.

Fingerprints View

- Displays the application's fingerprints library.
 - The fingerprint library is managed from here.
- Fingerprints classes and fingerprints can be added and deleted from this view.
 - This is where anything having to do with fingerprint management is done.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-28. Fingerprints View

Document Hierarchy View

- The document hierarchy describes the structure of the documents that your application is designed to process.
 - The levels within the hierarchy are batch, document, page, and field.
- It also displays how each element in the structure is related to the other elements and how each object is processed.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-29. Document Hierarchy View

Properties View

- Displays the properties for the selected document hierarchy object.
 - If the document hierarchy is locked for editing, existing properties can be modified.
 - In the properties panel, recognition options for the selected object can be specified.
- Since Datacap supports multiple recognition engines, the properties panel can display ICR/C, BAR/P, and OCR/S panels by default.
 - Other panels can be accessed by right-clicking within the Properties panel and selecting Show tabs.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-30. Properties View

Image View

- Displays the selected fingerprint image and any recognition zones.
 - New zones can be drawn or existing zones corrected or changed in the image view panel.
- If fingerprints were created by using full page recognition, the recognition results can be seen in the Text tab.

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-31. Image View

Exercise: Datacap Studio Zones

Requires:
Course Exercise Guide
Student system

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-32. Exercise: Datacap Studio Zones

Exercise objectives

- Explore the Datacap Studio – Zones view



Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-33. Exercise objectives

Unit summary

- After completing this unit, you should be able to:
- Describe the components in Datacap Studio for application development process
- Use the Datacap Studio

Datacap Studio Introduction

© Copyright IBM Corporation 2017

Figure 4-34. Unit summary

Unit 5. Add a Multi Page Document

Overview

This unit demonstrates adding a document with multiple pages to the document hierarchy and recognizing hand written input and mark input. It also demonstrates data validation with database lookup.

Objectives

- After completing this unit, you should be able to:
- Create a multi-page form application
- Configure ICR and OMR fields
- Implement Validation with database lookup

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacapoc.doc/datacap_9.0.1.htm

Unit Objectives

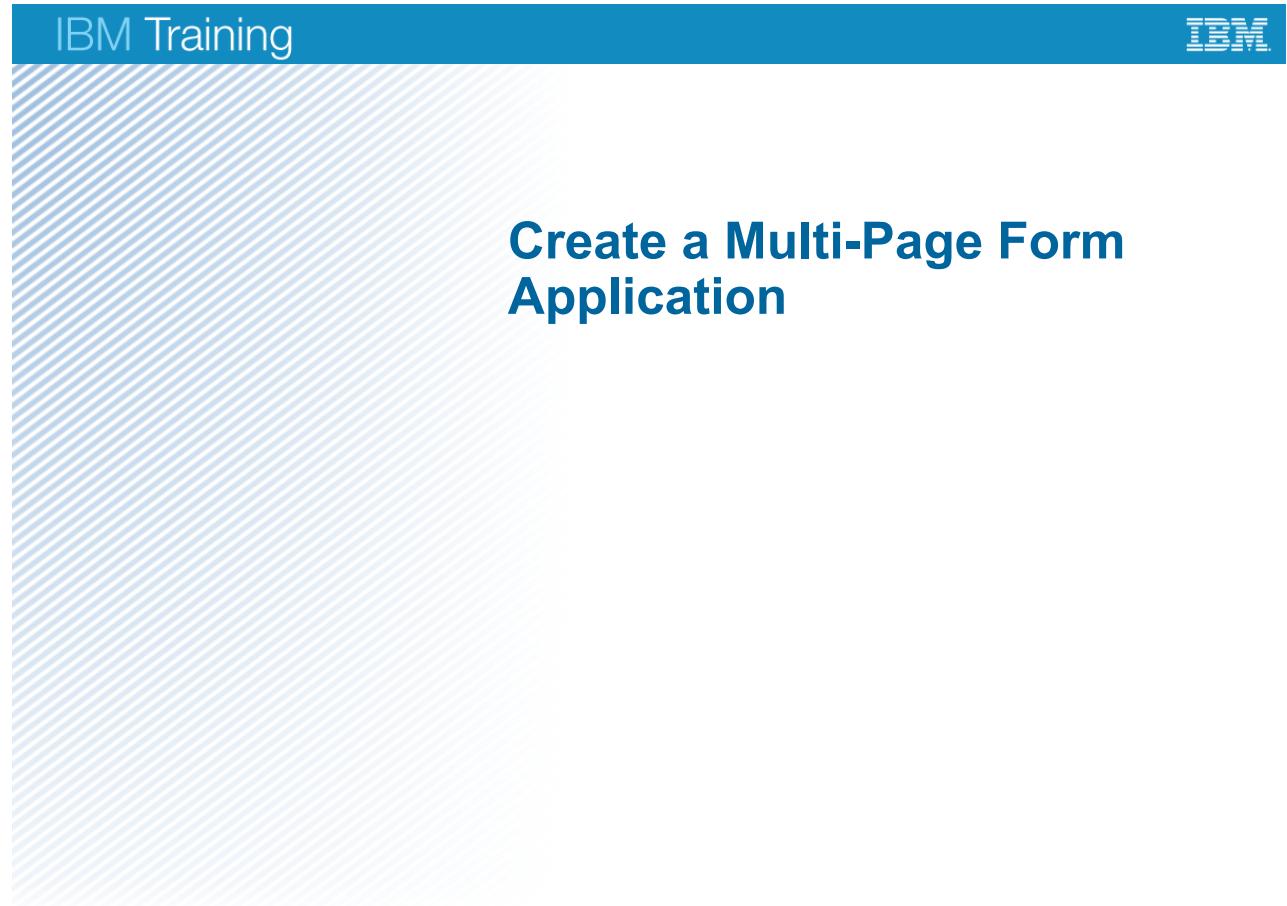
- After completing this unit, you should be able to:
- Create a multi-page form application
- Configure ICR and OMR fields
- Implement Validation with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-1. Unit Objectives

Lesson 5.1. Create a Multi-Page Form Application



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-2. Create a Multi-Page Form Application

Topics

Create a Multi-Page Form Application

- Recognize with ICR and OMR
- Validate with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-3. Topics

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users.
- To configure most Datacap capture applications, you must know how to configure a document with multiple pages. The page images that you use in the class are rigid forms. They are best processed with Form Template based application. Also important is configuring Recognition rulesets for ICR (Intelligent Character Recognition) for handwriting and OMR (Optical Mark Recognition) for check boxes.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-4. Why is this lesson important to you?



Image Samples

- Humanitarian Relief Fund

Humanitarian Relief Fund

This is a non profit organization.
This form is for Disaster Relief purposes only.

PLEASE CHOOSE YOUR PAYMENT PREFERENCE FROM ONE OF THE FOLLOWING OPTIONS	Barcode
<input checked="" type="checkbox"/> One Time <input type="checkbox"/> Monthly <input type="checkbox"/> Annual	PLEDGE AMOUNT Total - Frequency - Amount - Total 10050000 10050000 03102005
PREFERENCE	PAYMENT METHOD
<input checked="" type="checkbox"/> Bill Me <input type="checkbox"/> Check Enclosed	SECURITY CREDIT CARD TYPE CreditCard <input type="checkbox"/> Visa
Enter Human Credit Card, a charge will automatically be done at the Frequency selected above.	
AMEX CARD NUMBER (must have 6 digits) 4398209956814326 EXP DATE ON CARD 0308 Barbara Gill <small>Barbara Gill</small>	
PERSONAL INFORMATION Please Print the Following Information LAST NAME GILL FIRST NAME BARBARA ADDRESS 1 101 MAIN STREET ADDRESS 2 CITY TARRYTOWN STATE NY 10898 ZIP CODE SOCIAL SECURITY NUMBER 082-53-9100	
PLEASE SELECT ALL THAT APPLY <input type="checkbox"/> Victim Relief <input checked="" type="checkbox"/> Fund Raising <input type="checkbox"/> Awareness	
YOU MAY HAVE YOUR CONTRIBUTION RECOGNIZED ON: RED CROSS Barbara Gill <small>Barbara Gill</small>	

- Donation Receipt

Donation Receipt from Humanitarian Relief Fund

Questions? Contact the Humanitarian Relief Fund at 777-888-9999.

Donated by	Barbara Gill
Address	101 Main Street
City	Tarrytown
State/Province	NY
ZIP/Postal Code	10898
Phone	678-555-6789
Type of donation	CASH
Description	Credit Card donation
Value	\$1005.00

Thank you for your generosity. We appreciate your support!

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-5. Image Samples

The image on the slide illustrates sample Humanitarian Relief Fund form and a donation receipt.

Humanitarian relief Fund Form

- Structured layout
- Hand written text fields
- Red Boxes represent potential OMR fields

Donation Receipt

- Structured layout
- Machine printed txt

Application template choice

- The Humanitarian Relief Fund form document and Donation Receipt have a structured format.
 - They always have the same fields.
 - The fields do not move, they are at a consistent location.
 - Ideally suited for the Form Template application.
- Create the basic application.
 - From FastDoc Admin (local mode).
 - Run the Application Wizard.
 - Create an RRS application.
 - Application name: HRF_Form.
 - Select Form Template.

Figure 5-6. Application template choice



Configure document, pages and fields view

- FastDoc (Admin) Datacap Server mode:



- Configure the Batch Structure.
 - Add documents, pages, and fields.
- Configure Ruleset Properties.
 - Configure ruleset properties and Test profiles and rulesets.
- Configure Fingerprints.
 - Add fingerprint classes, add a fingerprint page, and map page fields to image zones.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-7. Configure document, pages and fields view

The screen capture shows the FastDoc “Configure documents, pages, and fields” window.



Add an HRF_Document and HRF_Page

- Add a document (HRF_Document) to the document hierarchy structure.
- Add a page (HRF_Page) to the HRF_Document.

Use rulesets from: ? Enable

- Add fields to the HRF_Page:
 - Zip
 - City
 - State
 - Frequency

	HRF_Page
□	Zip
□	City
□	State
□	Frequency

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-8. Add an HRF_Document and HRF_Page

The screen capture shows the options for copying page configuration from an existing page, when you are creating a new page.

In this example, HRF_Document, HRF_Page (first page), and four fields on the HRF_Page are used:

HRF_Document

|

 HRF_Page

 / / \ \

 Zip City State Frequency

Remember:

- When defining new Documents and Pages to always click Enable for the Use rulesets from?
 - If you are defining a new Document, select Document from the option list.
 - If you are defining a new Page, select Page from the option list.
- If you select the Document or page option, then the default rules from the template for the document or page are copied for your new object.
- If you do not enable the use of the default rulesets, then you are responsible for creating your own rulesets for managing the objects.
- Note; Click the? Next to Use rulesets from:? for helptext that describe this option.

HRF_Page Settings

- With HRF_Page selected and the Settings Tab selected:
 - Define Minimum, Maximum, and Order for the HRF_Page
 - Minimum 1
 - Maximum 1
 - Order 1

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-9. *HRF_Page Settings*

It also defines the Min, Max, and Order for the HRF_Page. All three variables have a value of 1

- There must be one, and only one, HRF_Page in an HRF_Document and it is always the first page.

Recognition considerations

- For the HRF_Page:
 - Configure recognition as full page zonal recognition.
- For the HRF_Page, Fields:
 - Text fields are hand written.
 - Recognition needs to be configured for ICR.
 - The Frequency, Payment Type, Credit Card Type, and Contribution Usage fields are optical marks fields.
 - Recognition needs to be configured for OMR.
- For Donation_Receipt pages, the fields are machine printed.
 - Recognition needs to be OCR.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-10. Recognition considerations

ICR = Intelligent Character Recognition (for hand writing)

OMR = Optical Mark Recognition

OCR = Optical Character Recognition (machine print text)

Data recognition is the stage during which you locate the fields that you want to capture and then convert the fields into character-based data.

- The data that is obtained from recognition is stored in the page data files that you set up in the document assembly stage.
- You can use several techniques that to identify pages.
 - Fingerprint matching: (The most widely used)

- Other non-fingerprint methods to analyze the page and identify and locate the fields are:
 - Text matching
 - Pattern matching
- If you used fingerprint match for page identification, you most likely used the fingerprint images to define the recognition zones.
- These zones are the fields that you want to read on each page.



Full page zonal recognition

- Page level recognition for HRF_Page and Donation_Receipt.
- In FastDoc, set these options at the Batch Structure Page level.



- For full page recognition, you get the field data directly from the full page recognition results.
- Otherwise, you need to run the recognition engine on each field zone to capture the data.

Figure 5-11. Full page zonal recognition

The image on the slide illustrates the Ruleset tab view of the of the FastDoc Configure, documents, pages, and fields window. The options shown are the options for full page zonal recognition.



Recognize hand written text fields (ICR)

- In FastDoc, set these options at the Batch Structure field level.

The screenshot shows the FastDoc Configure interface. On the left, there is a tree view of documents and fields: HRF_Form, Document, HRF_Document, HRF_Page, Zip (which is selected and highlighted with a red box), and City. On the right, there are three tabs: Settings, Ruleset, and Fingerprints. The Ruleset tab is active. Under the Ruleset tab, the 'Ruleset' dropdown is set to 'Recognize Pages and Fields'. Below it, there are two options: 'Read Field' (checkbox checked) and 'Add page recognition text to the zone' (radio button unselected). The 'Read hand print in zone' option (radio button selected) is highlighted with a red box.

- More ICR configuration is done in Datacap Studio (covered in next lesson).

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-12. Recognize hand written text fields (ICR)

The screen capture shows the Ruleset tab view of the of the FastDoc Configure, documents, pages, and fields window.

ICR = Intelligent Character Recognition (for hand writing)

For the HRF_Page > Fields

Read field check box

- It is a general option required for any recognition to be done on the field.

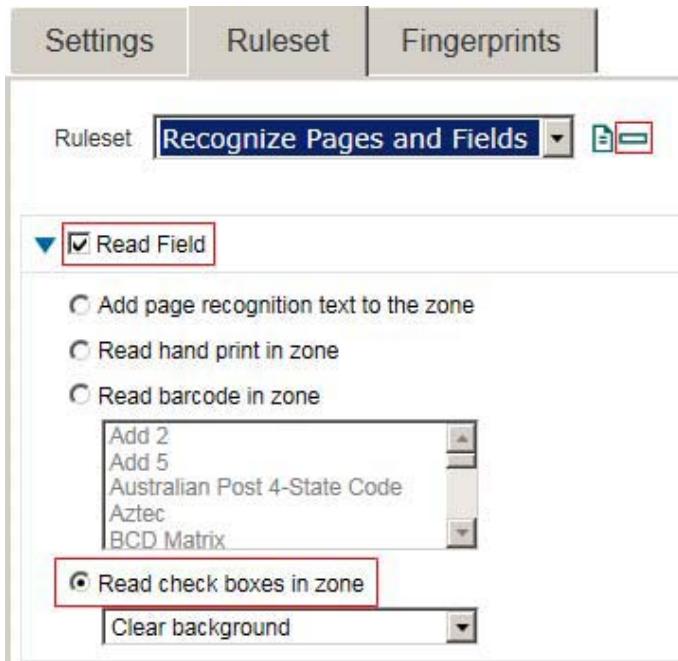
Read hand print in zone option

- If text fields are hand written, this option activates ICR field level actions in the “Recognize Pages and Fields” ruleset.
- More configuration is needed in the Datacap Studio Zones tab, in the ICR/C tab of the properties pane. (The Datacap Studio configuration is covered in the next lesson).



Recognize mark zones (OMR)

- For the HRF_Page > mark zone fields
- In FastDoc, set these options at the Batch Structure field level.



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-13. Recognize mark zones (OMR)

The screen capture shows the Ruleset tab view of the of the FastDoc Configure, documents, pages, and fields window. The options shown are the page recognition option.

OMR = Optical Mark Recognition

For the HRF_Page > Fields

These options are set for the zone that surrounds all of the mark zones. (In your example, it is the Frequency field.)

Read Field check box

- Is a general option required if any recognition it to be done on the field.

Read check boxes in zone option.

- If fields are check boxes, this option activates OMR field level actions in the “Recognize Pages and Fields” ruleset.
- All zones that are defined within this zone are interpreted as OMR zones.
- More configuration might be needed in the Datacap Studio Zones tab, in the OMR tab of the properties pane. (The Datacap Studio configuration is covered in the next lesson).
- More configuration is required on the Settings tab in FastDoc.



Field Settings tab

- Set Optical mark Enable.
- Configure Optical mark option list for the Verify step.

Value	Display Text
1	One Time
12	Monthly
4	Quarterly
2	Annually (this year & next)
1	

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-14. Field Settings tab

The screen capture shows the Settings tab view of the of the FastDoc Configure, documents, pages, and fields window.

For the HRF_Page > Fields

These options are set for the zone that surrounds all of the mark zones. (In your example, it is the Frequency field.)

Optical mark Enable check box

- Select to be able to run OMR recognition
- You must add a dictionary item for each check box.

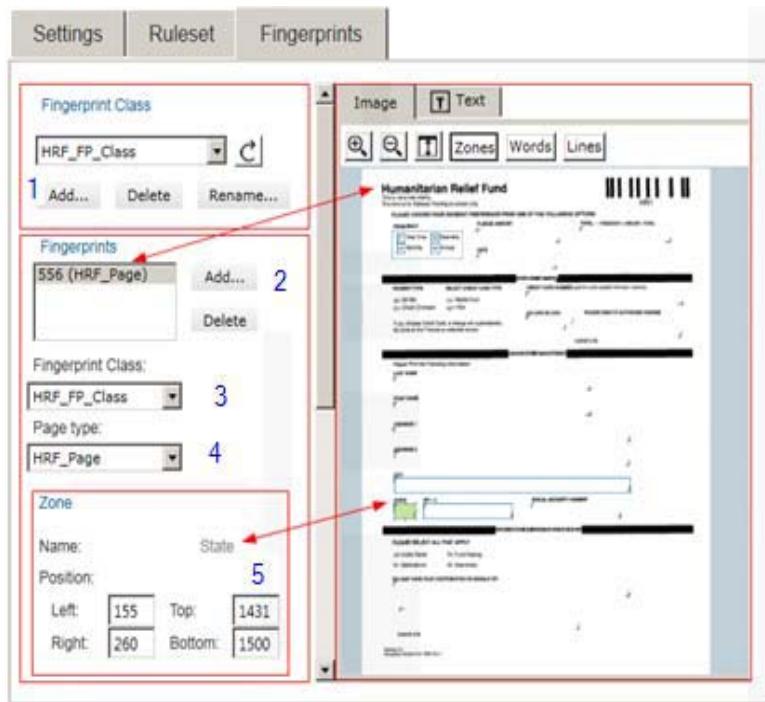
Add Dictionary items

- Building a dictionary table does two things
 - It adds an OMR field under the (Frequency) Check box zone field.
 - Creates an entry in an option list that can be used on the Verify panel if you need to override the OMR value for this field.

There is also an option to allow multiple selections from the list. "Allow multiple selections".

Configure Fingerprints

- Add a Fingerprint class
- Add a Fingerprint image
- Select a Fingerprint class
- Select page type
- Mark Zones



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-15. Configure Fingerprints

The screen capture shows the Fingerprints tab view of the of the FastDoc Configure, documents, pages, and fields window.

- Fingerprints are defined to enable the identification of images by the position of data items on a page.
- When an image is identified by matching it to a fingerprint, it is quick to extract the data values from the fields that the zonal information defines that is stored in the fingerprint.
- Locating data on an image with the zonal information stored in the fingerprint is more efficient than other methods of locating data like doing key word searches.
- You can assign the fingerprints that you create to fingerprint classes to differentiate the forms that you receive from various sources.
- Fingerprint classes can be used to group fingerprints based on the source of the form.

Add a Donation_Receipt Page

- Add a second page to the HRF_Document and name it Donation_Receipt.
- Add fields:
 - Donor_Name
 - Value
- Define Min, Max, and Order for
 - the Donation_Receipt:
 - Minimum 1
 - Maximum 0
 - Order 0

 Donation_Receipt
□ Donor_Name
□ Value

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-16. Add a Donation_Receipt Page

The screen capture shows the rule for identifying a second or subsequent Donation_Receipt. If you set Minimum to 1, then must always be at least one donation receipt included in a document. If you set Minimum to 0, then having no donation receipt is acceptable. Note: If you set Maximum to 1 and then do not include a donation receipt in your batch, then the batch is routed directly to FixUp to correct the problem before processing the batch.

Exercise: Create a Multi-Page Form Application

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-17. Exercise: Create a Multi-Page Form Application

Exercise introduction

- Create a Form Template based application
- Add a Second page to the document

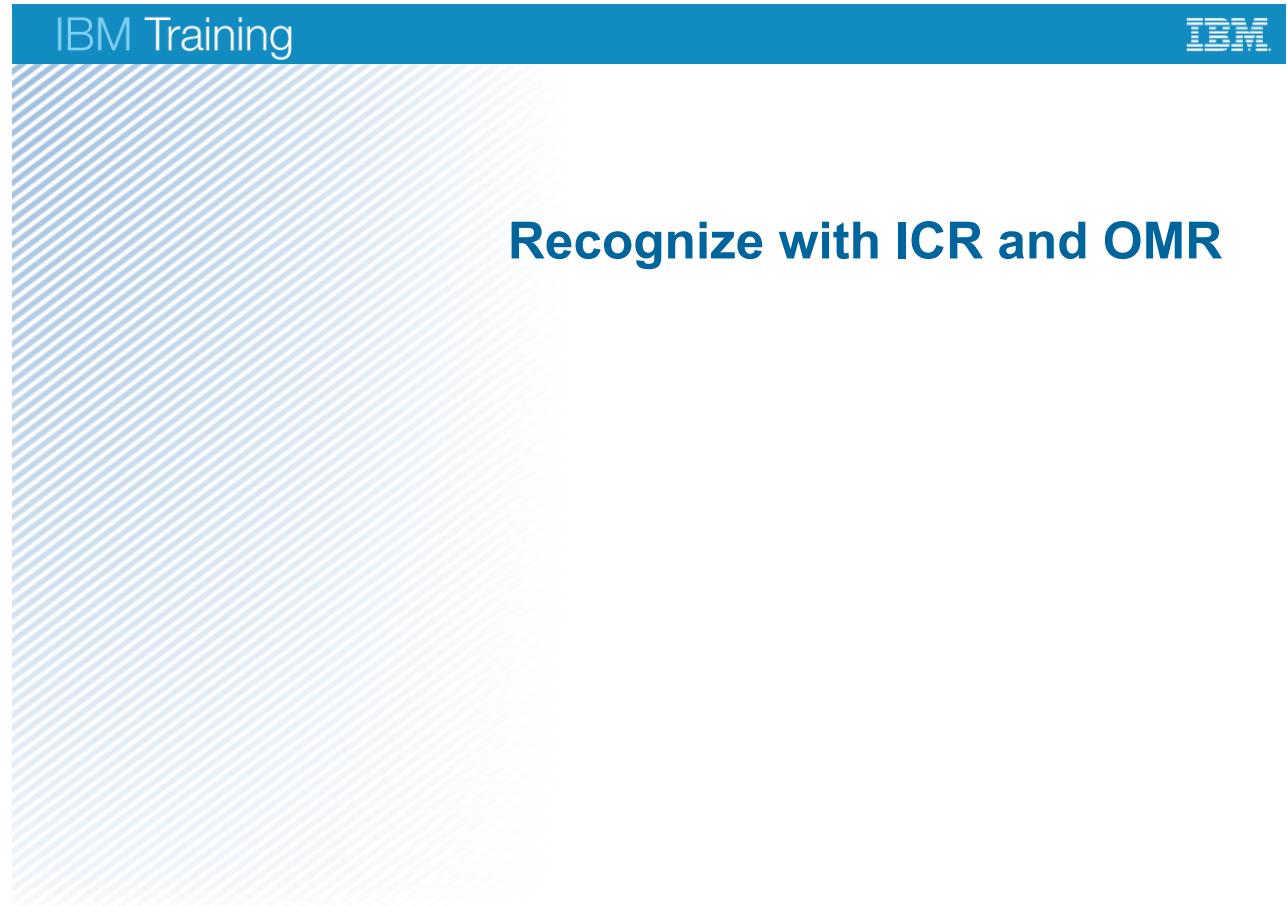


Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-18. Exercise introduction

Lesson 5.2. Recognize with ICR and OMR



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-19. Recognize with ICR and OMR

Topics

- Create a Multi-Page Form Application
- ▶ Recognize with ICR and OMR
- Validate with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-20. Topics

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users.
- To build a multi-page Datacap application, you must improve your ability to recognize pages and fields with more advanced techniques. In this lesson, you use Datacap Studio to expand on the configuration for hand written text fields recognition with ICR, and multiple option check boxes recognition with OMR.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-21. Why is this lesson important to you?

Overview

- OCR and ICR engines:
 - ocr_a: ABBYY FineReader OCR engine
 - OCR_s: Nuance (formerly ScanSoft) OmniPage OCR engine
 - ocr_sr: Newer implementation of the Nuance OmniPage OCR engine
 - icr_c: Open Text RecoStar ICR engine
- The Forms Template uses the following actions for recognition:
 - On a page that only does hand written text fields and optical mark fields:
 - Page recognition
Zones: ReadZones() or FPXML: ReadZonesFPX()
 - Field recognition
For hand writing icr_c: RecognizeFieldICR_C()
For optical mark ocr_a: RecognizeFieldOCR_A()
 - On a page with machine print:
 - Page recognition
Machine print ocr_sr: RecognizepageOCR_S()
 - Field recognition
Machine print Recog_Shared: SnapCCOtoDCO()

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-22. Overview

The RecognizePageOCR_S() is the action that is used for full-page recognition creates the .cco file that is later used for fingerprinting and locating data on a page.

- Datacap includes: three optical character recognition (OCR) engines,
 - OCR_a: ABBYY FineReader OCR engine
 - OCR_s: Nuance (formerly ScanSoft) OmniPage OCR engine
 - OCR_sr: Newer implementation of the Nuance OmniPage OCR engine
- Datacap includes: one Intelligent Character Recognition (ICR) engine that you can use to do full-page recognition:
 - ICR_c: Open Text RecoStar ICR engine
- The OCR engines work well with machine-printed text,
- The ICR engine works well with hand-printed and machine-printed text.
- The ICR_sr engine is not used to do full-page recognition in this application.

Recognize with the ICR Engine

- Uses the OpenText RecoStar engine to recognize constrained (unconnected) hand- or machine-printed characters.

Recognize Actions	Description
EnableLoggingICR_C	Enables or disables event log for the ICR/C engine.
RecognizeFieldICR_C	Does character recognition on the current field.
RecognizeFieldVoteICR_C	Does recognition on the current field zone and compares the result to the existing field value, character by character, raising the confidence level when the characters match and lowering it when they do not.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-23. Recognize with the ICR Engine

The image on the slide illustrates a table of Recognize actions with descriptions.

ICR recognition

The ICR, or Intelligent Character Recognition, engine is an advanced optical character recognition (OCR) engine. It uses the OpenText® RecoStarTM engine to recognize constrained (unconnected) hand- or machine-printed characters.

The length of a field is important for ICR. If the field length is set to 9, the ICR engine divides a zone into nine pieces and looks for a character in each piece.

More Recognize Actions

Recognize Actions	Description
RecognizePageFields2CCO_I CR_C	Does recognition for all zoned fields on the page.
RecognizePageFieldsICR_C	Does recognition on all fields that are configured for ICR/C.
RecognizePageICR_C	Does full page recognition with the ICR/C engine.
RecognizePageToPDFICR_C	Does full page recognition and stores the results in a PDF file.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-24. More Recognize Actions

The image on the slide illustrates a table of Recognize actions with descriptions.

IBM Training

Datacap Studio Zones tab

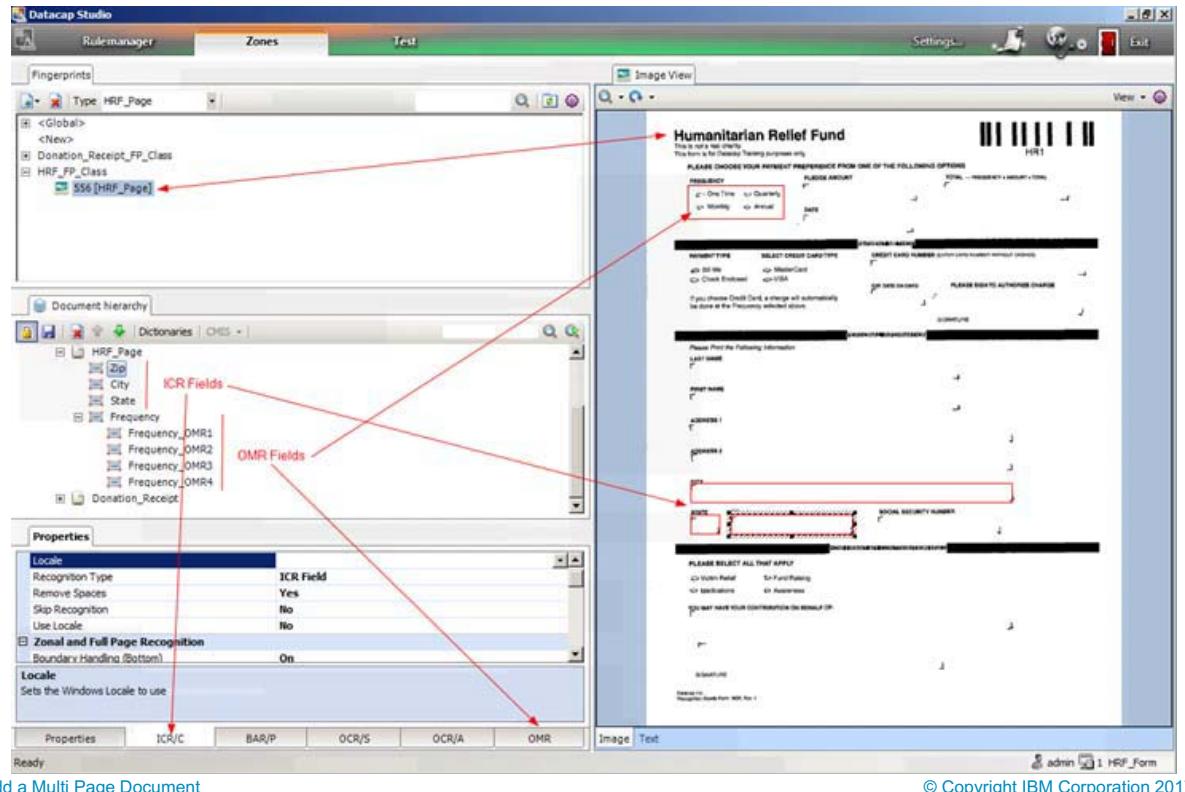


Figure 5-25. Datacap Studio Zones tab

The screen capture on the slide shows a view of the Datacap Studio Zones tab.

- The previous lesson stated that more configuration for the ICR fields is required in Datacap Studio.

ICR configuration

- Verify that zones are correctly defined for each ICR field.
 - Click each ICR field that is defined in the Document hierarchy pane and verify that the correct field in the Image View gets highlighted.
Note: you must lock the document hierarchy view before the zones become visible in the Image View.
 - Click each ICR field in turn in the Document Hierarchy pane.
 - Click the ICR/C tab in the properties pan in the lower left corner.
 - Configure the ICR properties as shown on the “ICR/C tab setting” slide.

OMR Configuration

- Verify that the OMR parent area is zoned. (Click the Frequency field in the Document Hierarchy pane and verify that the zone below the Frequency heading is marked around the four optical mark fields.
 - Notice that the four OMR fields Frequency_OMRx are defined in the Document Hierarchy field but there are no zones that are defined for them.
 - These fields were created when you defined the Frequency Dictionary table in FastDoc.
- Zone the four OMR fields.
 - Zoom in on the Frequency zone so that it is large enough to properly zone the OMR fields.
 - Make the Zones rectangles of the same size and space line them up vertically and horizontally.
 - Remember that you must define the zones in the same order that you used when you created the Dictionary table.
 - Review the OMR properties as shown on the “Select OMR Properties View” slide.

ICR/C Tab Settings

- Recognition Type
- Boundary handling
- Character set
- Country
- Delete intermediate project file
- Font
- Length
- Logical context
- Number of lines
- Pattern
- Pitch
- Reader
- Syntax
- Triagram Mode

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-26. ICR/C Tab Settings

ICR field setup

ICR settings for a field must be done on the ICR/C tab. Ensure that some white space is left around the boundaries while marking a field in fingerprints.

Key settings on the ICR tab:

- **Local:** Warning. Clear this field.
- **Recognition Type:** Select ICR for ICR fields.
- **Boundary handling:** Turn on boundary handling, especially if there is not much space around each field. The application reads and extracts data from the outer parts of marked zones if these fields are on.
- **Character set:** Select the appropriate option if you want to extract specific data from the field. For example, select 0-9 for ZIP code.
- **Country:** Select the country or language of the data to be extracted.
- **Delete intermediate project file:** Set this value to False if you want the application to generate a log file for debugging. The file is saved in your Temp folder with a .rsp extension.
- **Font:** Select handprint for ICR.

- **Length:** Enter the number of parts that you want the zone to be divided into before extracting the data. For example, if you are identifying the ZIP code field, enter 9 here. The application divides the ZIP code into nine sections and looks for data in each field.
- **Logical context:** Set this value to **On** to set the application to check for the most likely value of a questionable character. For example, if you set this value to **On** for a ZIP code, the application identifies the character "L" or "l" as "1".
- **Number of lines:** Set this value to **Unknown** if the number of lines in the zone is unknown.
- **Pattern:** Use this option to match the specified RegEx (regular expression) pattern.
- **Pitch:** Select **Fixed** if the values in the zone are of the same size. If the values vary, select **Variable**. If you are not sure, select **Unknown**.
- **Reader:** Use this option to assign either the RecoStar or AEG reader to read the data. Select the Voter option here. This option is a combination of RecoStar and AEG.
- **Syntax:** Indicates the most likely data type in the zone. For example: date, numeric, text.
- **Triagram Mode:** Set this value to On. The application checks if the values next to the recognized value are any indication of the most likely value of the character.

Recognize OMR

- Optical mark recognition (OMR)
 - Used to electronically extract intended data from marked fields on printed forms.
 - Scans a printed form and reads predefined positions and records where marks are made on the form.
- OMR methods:
 - Pixel threshold evaluation method:
 - Is more reliable for dropout check boxes.
 - Can also be used to read fill-in bubbles on a response form.
 - Is difficult to set up.
 - OCR/A check box recognition method:
 - Is easy to set up.
 - Works well with non-dropout check boxes.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-27. Recognize OMR

OMR fields

Optical Mark Recognition (OMR) is a technology that is used to electronically extract intended data from marked fields on printed forms. These marks are check boxes and complete fields like bubbles. The OMR technology scans a printed form and reads predefined positions and records where marks are made on the form.

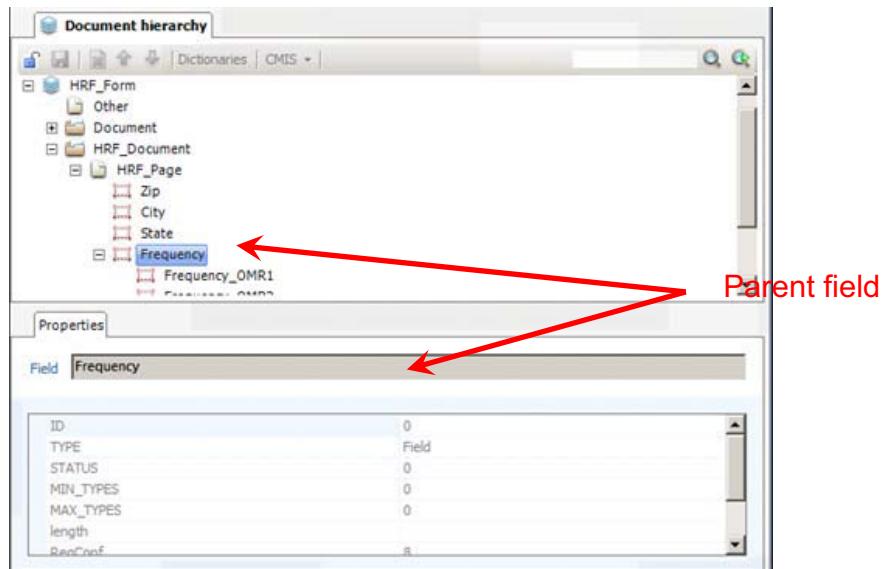
OMR field setup

Datacap employs optical mark recognition (OMR) to determine whether a check box option is selected or not. There are two basic techniques:

- OCR/A check box recognition method
- Pixel threshold evaluation method

IBM Training

Establishing OMR Parent fields



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-28. Establishing OMR Parent fields

The image on the slide illustrates the Document hierarchy and Properties panes and highlighting the OCR parent name filed.

Establishing Parent fields

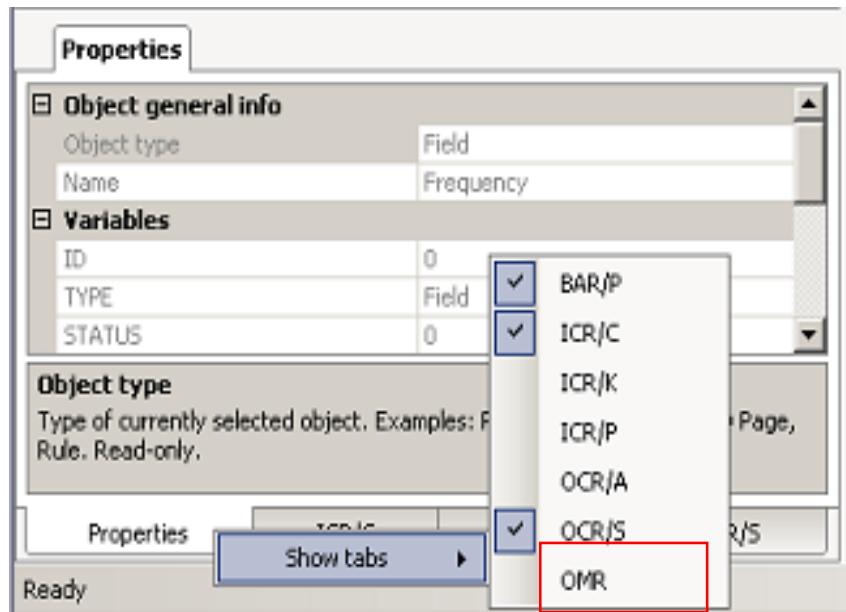
Processing pages with check box options:

- The system automatically defines the check box options as subfields of the parent Frequency field within the document hierarchy.
- You must outline the subfields and the parent field when drawing the recognition zones.

When you define the recognition zones, you need to define the positions of the parent fields and the subfields.

Select OMR Properties View

- For pixel threshold evaluation method



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-29. Select OMR Properties View

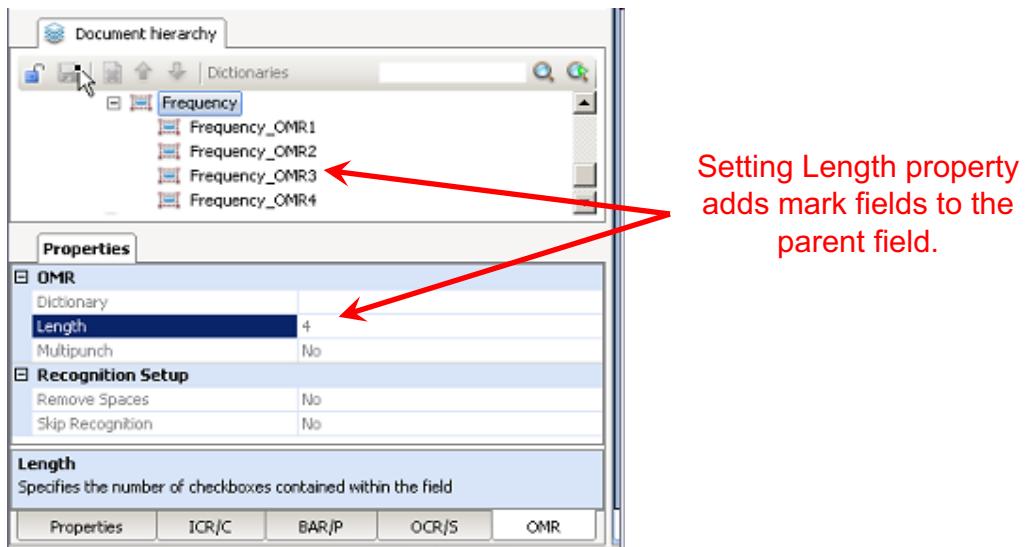
The image on the slide shows the selection of the OMR option from the properties pane.

You must configure the OMR settings for specific zones with the OMR tab in the Properties pane on the Datacap Studio Zones view. The OMR tab is not shown by default, so you must enable it, as described in the following section.

Right-click any existing tab and click **Show tabs**. Then, select the **OMR** option.

IBM Training

Set OMR Length Property



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-30. Set OMR Length Property

The image on the slide illustrates Document hierarchy tab showing the Frequency OMR fields and the corresponding Length field in the Properties pane the represents the number of OMR zones.

With the Pixel threshold evaluation method

Click the OMR tab to show the settings that the OMR recognition engine uses when doing recognition on the selected field.

The Length property defines how many mark fields are defined for the parent field.

Note: In this class, you define the Frequency field in FastDoc. When you create the OMR dictionary list for manually selecting the OMR field options at Verify time, the subfield corresponding to the list options are automatically defined. All that is left to be done is to associate the mark fields with the zone on the image where the data is read from.

The OMR Settings

Check mark type	Select “Square background” to read non-dropout Check boxes. This setting is stored in the document hierarchy with the OMRTyp variable, where 0 is “Square background”: <code><V n="OMRTyp">0</V></code>
Length	This setting reflects the number of OMR subfields and is set automatically.
Multipunch	This setting is the same as the MultiPunch variable where 1 is “Yes”: <code><V n="MultiPunch">1</V></code>

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-31. The OMR Settings

The image on the slide illustrates the OMR setting options.

With the Pixel Threshold Evaluation Method

The pixel threshold evaluation method uses the RecogOMRThreshold action in the “Recog_Shared” library and is the method that is used in an exercise.

Specifying the Threshold and Background levels

The RecogOMRThreshold action takes two parameters:

- Threshold: Specifies the percentage of black pixels, which cause the option to be considered selected.
- Background: Used to determine the confidence level and specifies the percentage that can be attributed to the check box outline plus any scanner noise:

Any zone with below the background value is considered not selected with high confidence. Any zone with between the background value and the threshold value is considered not selected with low confidence.

Any zone with greater than $(2 * \text{Threshold} - \text{Background})$ is considered selected with high confidence. Any zone with between Threshold and $(2 * \text{Threshold} - \text{Background})$ is considered selected with low confidence.

Exercise: Recognize with ICR and OMR

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-32. Exercise: Recognize with ICR and OMR

Exercise introduction

- Configure ICR and OCR fields

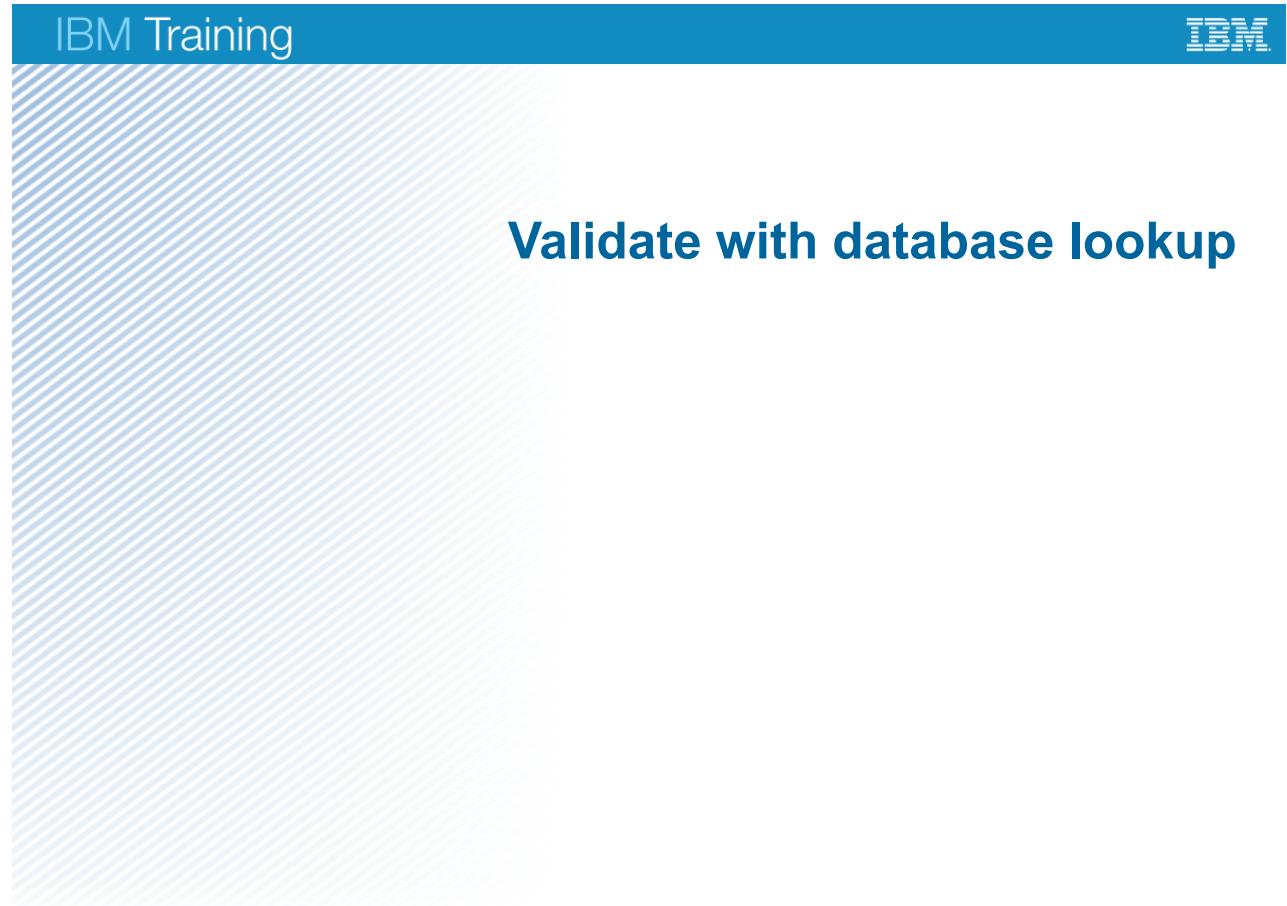


Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-33. Exercise introduction

Lesson 5.3. Validate with database lookup



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-34. Validate with database lookup

Topics

- Create a Multi-Page Form Application
- Recognize with ICR and OMR
- Validate with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-35. Topics

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users.
- To build intelligent Datacap applications, you must be expert at doing database lookups to validate data that is interpreted on the captured pages. In this lesson, you validate a postal code with database lookup to illustrate this important skill.

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-36. Why is this lesson important to you?

Steps to Configure Lookup

- Configure a connection string.
- Validate the new database connection.
- Configure Validate Fields Ruleset for field validation.
- Configure Lookup parameters in the Validate Fields Ruleset.
- Save your edits.

[Add a Multi Page Document](#)

© Copyright IBM Corporation 2017

Figure 5-37. Steps to Configure Lookup

Overview

The purpose of validation is to determine whether captured data conforms to specified business rules. For example, you need to determine whether the vendor information matches the information that is stored in a database of approved vendors.

Use database lookup to validate a ZIP Code

In this section, you use the Zip field on an existing page to demonstrate how to test whether a field contains a value that is in a list of allowed values. For each Zip field that the ICR process detected, you test the field value.

Configure a connection string

- Use Datacap Application Manager to connect to the lookup database.
- Copy the lookup database into the application root folder.

Sample: C:\Datacap\Expense

- Click the ellipsis at the end of the database string to open the ‘Data Link Properties’ window to configure the link to the database.
- Select the ‘Database type and properties name’: **Microsoft Access (Jet)**

Browse to the database: **C:\Datacap\Expense\Zipac.mdb**

Validate the new database connection

- Click test connection and check that you get Success status.

Configure Validate Fields Ruleset for field validation

- Use FastDoc to open the application.
- Select the index field for the lookup database.
- Configure validation parameters for the field.

Configure Lookup parameters in the Validate Fields Ruleset

- Click the Lookup check box and configure the lookup parameters.
- Select the database.
 - Type the SQL statement to do the lookup.
Sample for the Zip field:
Select Zip From Zipac Where Zip='+@p\Zip+';

Save your edits

- Click Save and close the FastDoc window.



Configure a lookup database connection string

- Configure a connection string.
- In the Application Manager.
- Click Lookup database
- Select the Database type.
- Browse to the database.
- Test the connection.



Add a Multi Page Document

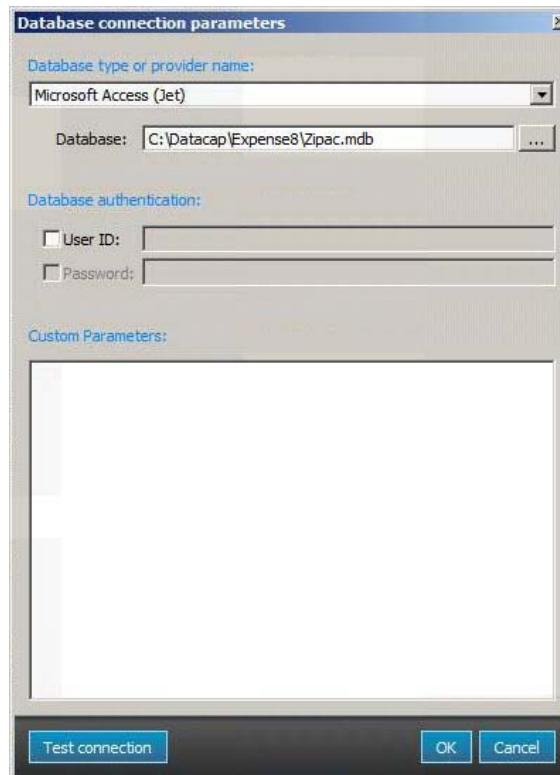


Figure 5-38. Configure a lookup database connection string

The screen captures show:

- The Database connection parameter window where you create the connection to the lookup database.
- I shows the Success message window after you have done the test connection operation.

Configure a connection string

Create a file on the desktop named **zipcode.udl** and double-click the file to open the Data Link Properties window to configure the link to the database.

Provider option: **Microsoft Jet 4.0 OLE DB Provider**

Database location: **C:\Datacap\Expense\images\HR\Zipac.mdb**

Look for the connection string in desktop\zipcode.udl.

Update the Lookup database string

Start Datacap Application Manager:

Start > All Programs > IBM Datacap Services > Datacap Application Manager

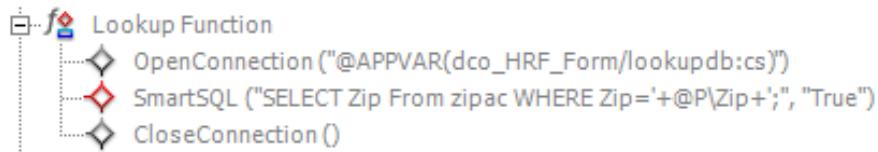
Update the Lookup database string on the Application tab.

Database Lookup Actions

- The Lookup library includes actions for connecting to external data sources and running SQL statements.

Lookup Actions	Description
OpenConnection	Uses a data source name or connection string to open a connection to a database.
SmartSQL	Run a SQL statement with support for smart parameters.
CloseConnection	Closes an open database connection.

- Validate Field > Lookup function:



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-39. Database Lookup Actions

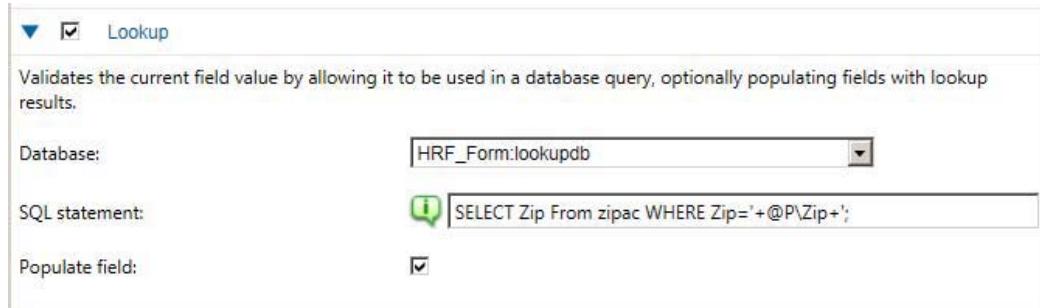
The Lookup Function shown in the screen capture shows one of the default functions activated when the Lookup option is selected in the Validate Fields ruleset.

- When you are using one of the Application templates (Learning or Form) for your application, the template automatically creates the necessary actions in the Standard Validate field ruleset. It also provides the plumbing for connecting to the Batch Structure objects and the Task profiles.
- These Lookup library actions that are shown in the slide table, are used in the template Validate Field Ruleset when the lookup option is selected in the Validate Fields ruleset in the FasDoc interface. They connect to external data source, run SQL statements, disconnect from the database, and populate the field.

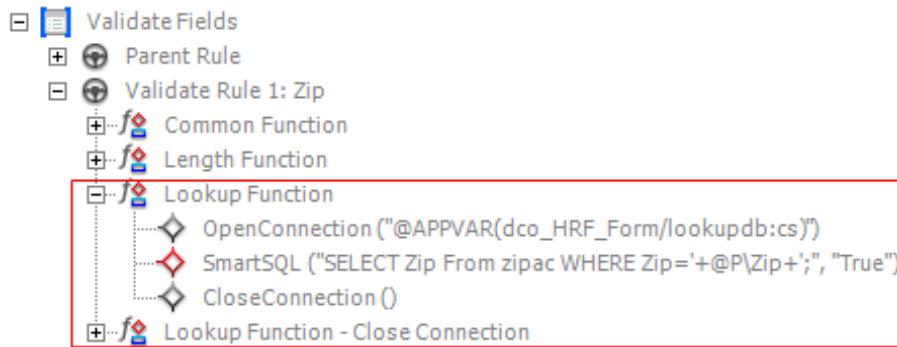
IBM Training

Validate Fields Zip Lookup Function

- FastDoc Lookup Configuration:



- Standard Template Validate Fields Ruleset:



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-40. Validate Fields Zip Lookup Function

The screen capture images on the slide show:

1. The UI from FastDoc for the Validate Fields Ruleset and the Lookup validate option.
 - Notice that the Lookup check box is set.
 - The configured parameters define:
 - The Database to read from: HRF_FORM: lookupdb
 - The SQL statement to read the zip value from the database with the zip field value as an index into the table.
 - The option to populate the Select field
2. The Ruleset that is created based on the FastDoc options.
 - Notice that the Open Connection action and SmartSQL action parameters match the values that are entered in FastDoc.

Exercise: Validate with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-41. Exercise: Validate with database lookup

Exercise introduction

- Configure a zip code database lookup
- Configure a zip code truncate ruleset



Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-42. Exercise introduction

Unit summary

- After completing this unit, you should be able to:
- Create a multi-page form application
- Configure ICR and OMR fields
- Implement Validation with database lookup

Add a Multi Page Document

© Copyright IBM Corporation 2017

Figure 5-43. Unit summary

Unit 6. Basic Learning Application

Estimated time

07:00

Overview

This unit introduces you to the process for building a basic Datacap learning application. It then walks you through the steps of refining the default image capture process as you make changes at each step of the process.

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/datacap_9.0.1.htm

Unit Objectives

- After completing this unit, you should be able to:
- Create a basic learning application with the Datacap Application wizard
- Apply image enhancement techniques and field recognition potential
- Define a document structure
- Create and expand Locate rules
- Create and expand Validate rules
- Export the scanned documents to an IBM FileNet repository

Lesson 6.1. Create a Learning Template application



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-2. Create a Learning Template application

Lessons

-  Create a Learning Template application
 - Configure a Learning Template application
 - Create locate rules
 - Create validate rules
 - Expand the locate rules
 - Expand the validate rules
 - Export to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-3. Lessons

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users.
- This lesson gives you the ability to decide when to use a Learning template rather than a Form template when you create your application. It also gives you the ability to create the learning application by using the application wizard and to do the basic application configuration.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-4. Why is this lesson important to you?

Create a Learning Template-based application

- Use the Application Wizard in FastDoc to create an application by using the Learning Template.
 - Name the application 'Expense'
 - Select the Learning Template when prompted to select the Application Template
 - Verify that no errors or warnings are encountered

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-5. Create a Learning Template-based application

Create new Application with Application Wizard

The Application Wizard can also be started from Datacap Studio.

The Application Wizard Gives you a head start on application development by generating a basic application framework.

- Copy an application – What do you get with this option?
 - A copy of a completely functioning application as a starting point.
 - You can modify and customize it to meet the functional requirements for your new project.
- Create an RRS application – What do you get with this option?
 - An option of selecting one of two built-in application templates as a starting point for your new application.
 - A complete application folder structure.
 - A skeleton Document hierarchy (DCO), which you can expand to provide the document structure for your planned document batches.
 - A functional set of built-in compiled rulesets.
 - A set of workflow task profiles, preconfigured to run data capture tasks.

- A custom mapping configuration to map rules from the rulesets to the DCO objects.

Note:

1. You can create and save your own custom templates to use for future projects.
2. When creating new DCO objects, you can configure them to inherit rulesets from existing objects of the same type.

Characteristics of a Learning Template application

- A Learning Template-based application is inherently different from a Form-based application.
- It processes documents entirely differently
 - It has both different setup and processing
- It is called a Learning Template because the application learns how and when to create fingerprints and does it automatically.
 - You might add a few fingerprints when developing your application, but after that, leave it to the built-in behavior of the Learning Template.
 - It is not recommended to explicitly configure fingerprints

Figure 6-6. Characteristics of a Learning Template application

Learning Template considerations

The first thing that you do when developing a new Datacap application is to decide which of the two built-in templates you should use. It isn't a difficult decision! A forms-based document type is one on which you know the types of data that you want to capture and where that data is on each image. Your documents are structured and consistent.

Characteristics of a Learning Template DocType

- Characteristics of a Learning Template document type.
 - The images are unstructured.
 - The types of data to capture are known.
 - The location of that data isn't known because it is different on each image.
 - For example, if you want to capture the date, amount and tax from different hotels, the receipt images for each hotel type are unique.
 - The data location to capture differs for each hotel receipt image, so fingerprints are not reliable at the start of building the application.

Figure 6-7. Characteristics of a Learning Template DocType

Document Type Considerations

Documents that are appropriate for a Learning Template-based application are unstructured. Documents that are made up of images which contain data that you want to capture but that data location isn't consistent. The location of the data might be different on every image.

Learning Template configuration

- The learning template uses generic document types, page types, and field types that are provided with the template.
 - It is important that you leave everything as it is except for the field objects that you define.
- There are four pages in the generic document object.
 - Main_Page
 - Trailing_Page
 - Attachment_Separator
 - Attachment
- It is important that you never change the Main_Page name.
 - Other page names can be changed although it's probably best not to do so!

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-8. Learning Template configuration

Learning Template Configuration.

- The learning template configuration uses generic document and page types that are provided with the template.
- The template contains four pages as the content of the generic document object:
 - Main_Page
 - Trailing_Page
 - Attachment_Separator
 - Attachment
- It is critical that you do not change those application objects – especially the Main_Page!
- All content must be put on the Main_Page

Main_Page

- Because you are developing a Learning Template application, you leave the name of the first page Main_Page.
- Main_Page represents two actual page types in the application you develop.
 - Rental_Agreement page
 - Air_Receipt page
- The only application objects that have custom names are fields.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-9. Main_Page

Main_Page and your application

During the processing of your documents, the Main_Page page object will represent either the Rental_Agreement page type or the Air_Receipt page type. However, your application process only the pages of your document that uses the Main_Page as the representation of one of those application page types.

Document integrity Values

- It is important, when you define a new page of a document, that you define its document integrity values.
- Each page is part of a document object.
- Document integrity answers the questions:
 - Is there a minimum number of this page type that must be in the document?
 - A maximum number? Must it always be the first page?
 - Can there be a variable number of this particular page type?
- You must configure the parameters that specify answers to those questions.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-10. Document integrity Values

Document Integrity Overview

The Document Integrity ruleset that is generated by the Application Wizard is configured to identify document integrity problems and send the batch to the FixUp task if required.

Document integrity

- It is critical to define document integrity correctly.
 - It is how document integrity problems are found so that a document with integrity problems can be sent to Fixup.
- The three parameters that define document integrity for each page of a document are Minimum, Maximum, and Order.
- In the application, you are developing, each document has one, and only one, Rental_Agreement OR Air_Receipt page.
 - It is always the first page if some of your documents are multi-page documents.
 - Therefore, all three parameters Minimum, Maximum and Order each has a value of one.
 - These three parameters are on the Page Settings tab on the FastDoc 'Configure documents, pages, and fields' view.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-11. Document integrity

Defining Document Integrity

Three Document Integrity parameters that must be defined:

Minimum – the minimum number of a particular page type in a document type

Maximum – the maximum number of a particular page in a document type

Order – where this page must be in the document object

In the Learning Template labs, you set Minimum, Maximum, and Order to have the value of 1. In other words, there must be 1 and only 1, Rental_Agreement or Air_Receipt page, and that page is always the first page in the document.

Learning Template fields (1)

- Fields are the only application content that isn't entirely generic in a Learning Template-based application.
 - It is important that you leave batch structure objects rules and rule associations as they are, except for the field objects that you define.
- You define four generic fields on the Main Page
 - Name
 - Date
 - Total
 - Reference_Number

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-12. Learning Template fields (1)

Generic fields

- Generic fields represent fields on one or more page types that aren't consistently in a particular location on a page.
- The fields are at different locations on different page types.
- The actual backend 'name' of the field on either page is not the generic field name at the point where you commit your document to that backend.
- It is critical to know the relationship between the processing name of a field object and the field's name on the backend to which you commit each document.

Learning Template fields (2)

- The four generic fields you define on the Main Page represent the following fields on the two page types

GENERIC FIELD NAME	Rental_Agreement Page	Air_Receipt Page
Name	Customer Name	Passenger Information Psgr
Date	Pickup Date/Time:	Issue Date:
Total	Total Charges	Total:
Reference_Number	Rental Agreement Number	TICKET #:

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-13. Learning Template fields (2)

Generic fields and how they relate to fields on the page of a document in a particular document class.

You have four generic field names that represent actual fields with different names in two different document classes.

This chart shows the relationship between the generic field name, “Name”, and the actual field name on two different page types.

For example, Name represents “Customer Name” on the Rental_Agreement page type and represents “Passenger Information” or “Psgr” on the Air_Receipt page type.

Exercise: Create a Learning Template application

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-14. Exercise: Create a Learning Template application

Exercise objectives

- Create a Learning Template application

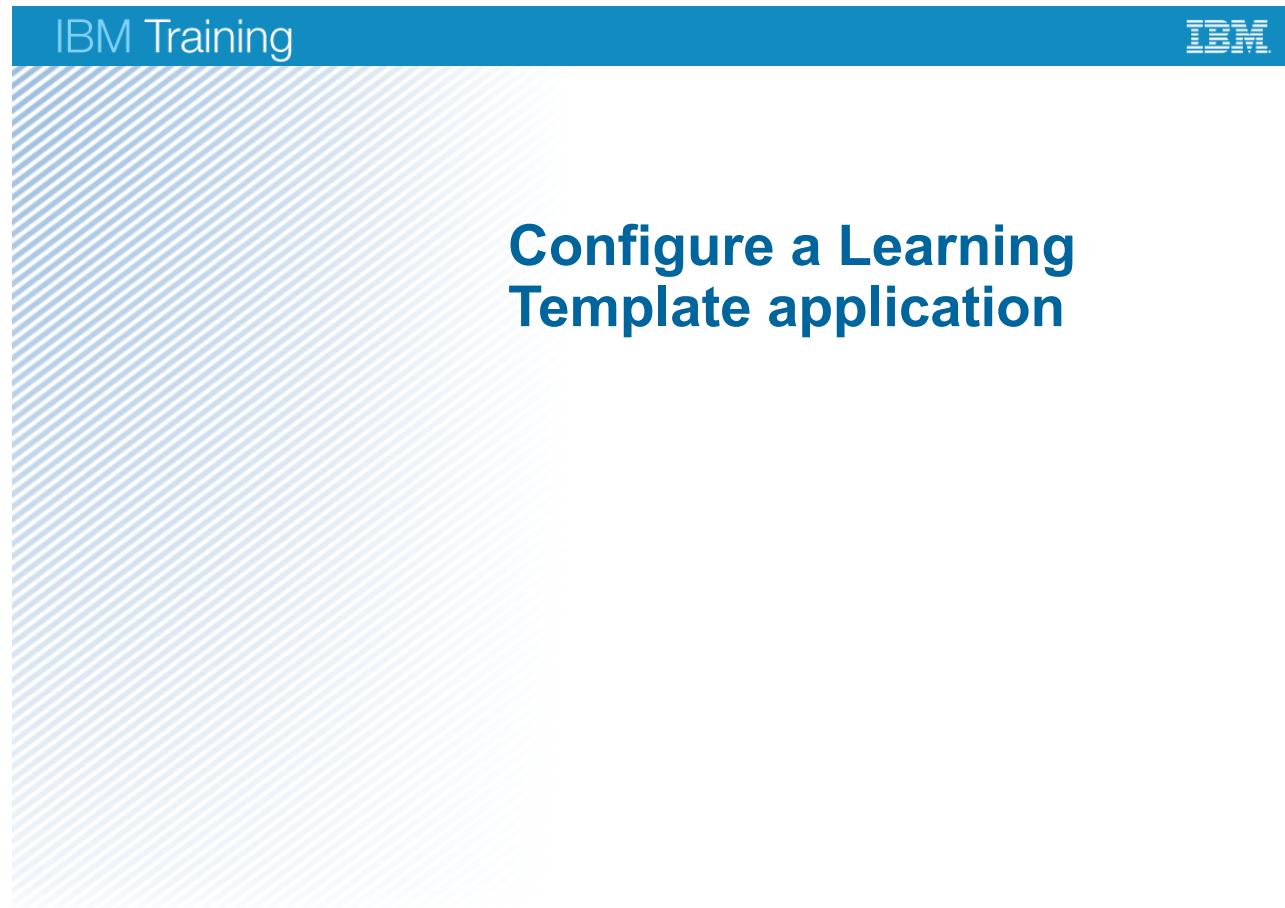


Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-15. Exercise objectives

Lesson 6.2. Configure a Learning Template application



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-16. Configure a Learning Template application

Lessons

- Create a Learning Template application
- ▶ Configure a Learning Template application
- Create locate rules
- Create validate rules
- Expand the locate rules
- Expand the validate rules
- Export to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-17. Lessons

Why is this lesson important to you?

- As a Datacap Solution/Application Builder, you build and deploy applications with the Datacap capture system and communicate solution details to the solution architect, administrator, and business analysts.
- When creating new applications, it is good practice to use the provided templates as a starting point. The Learning Template is used when processing images that differ in format but contain the same content to extract for indexing.

Figure 6-18. Why is this lesson important to you?

The Learning Template and fingerprints

- There is no need to set up fingerprints with this template.
 - If you do need to setup fingerprints, do so at the start of building the application by clicking and keying, for example:
 - You should not be generating fingerprints beyond the initial application creation and early configuration!
- Fingerprints are automatically generated as a user inputs data for the various document types and versions.

The Learning Template workflow

- The Learning Template application sets up a workflow.
 - Rules like Locate rules are used to teach Datacap about the different page type formats as they are encountered.
 - Fingerprints are automatically generated when page types are determined.
- For images where data isn't found, the verifier is prompted to click the image and identify where that data is located.
 - This Click'N'Key process populates the data into the data set so that the Learning application can automatically find the data the next time that type of image is encountered.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-20. The Learning Template workflow

The following procedure is the workflow process that the **Learning Template** uses for multi-page variable sized documents that use the **DemoMultiFormat** job.

1. Start FastDoc in Datacap Server mode. 2.
2. Log in to the **Learning** application.
3. Click the VScan shortcut. This task runs a disk scan on the image files that are stored in the C:\Datacap\appname\images\input folder.
 - a. Select the DemoMultiFormat job.
 - b. Runs the **Import**, **Convert Images**, and **DeleteSourceImages** rulesets. FastDoc returns to the **Shortcuts** screen.
4. Click the Profiler shortcut. This task sets up the form by determining where the data that you want to process is on the image.
 - a. Runs the Managed Rotation ruleset to enable rotation of the image if required.
 - b. Runs the Image Enhancement ruleset to clean up the imperfections in the image. Select Auto Rotate to enable the Managed Rotation ruleset.

- c. Runs the PageID ruleset to identify the path on which the document came in and creates the document structure. If the path is the same as another document, assumes that it is the same document type.
 - d. Puts all of the data on the Main_Page page type and treats every image as its own document. When the ruleset gets to a Separator sheet, it treats the next page as Main_Page. The pages after that are treated as a Trailing_Pages until the next Separator sheet is encountered.
 - Separator sheets are the method for grouping multiple pages into a document when using the Learning Template.
 - e. Runs the Recognize ruleset to identify the page types as Main_Page and Trailing_Page.
 - f. Inserts a Doc_Separator page in front of the Trailing_Pages..
 - g. Inserts an Attachment_Separator page in front of the attachments. Recognition is not run on the attachments.
 - h. Runs the Fingerprint ruleset to build a fingerprint with a fingerprint ID for each document type.
 - i. Runs the Locate ruleset to search CCO for data, such as Social Security Number, Invoice Number, Date.
 - j. Runs the Lookup ruleset to take the fingerprint ID and find the fingerprint class for it.
 - k. Runs the Validate ruleset on the fields to validate the data in every field and make corrections where needed.
 - l. Runs the Routing ruleset to route the forms that passed validation to the export task.
5. Click the Verify shortcut.
 - a. Runs the Validate ruleset on the form to remember the field information and automatically run it the next time that it is recognized.
 6. Click the Export shortcut to run the ruleset that you configured for this job.
 - a. Runs the SetStatuses ruleset to set up the exporting status of the document.
 - b. Runs the PreExport ruleset to use the specified export settings for the document.
 - c. Runs the Export ruleset to write the file to a specified location on the disk.
 - d. Runs the Export to FileNet Content Manager ruleset to store the file in the IBM® FileNet® Content Manager repository.
 - e. Runs the Export to IBM Content Manager ruleset to store the file in the IBM Content Manager repository.
 - f. Runs the ProcessExceptions ruleset to handle exceptions to regular processing. For example, what to do with rescans and unwanted documents.

More about the Learning Template workflow

- There is no separate PageID task.
 - It is part of profiler because the assumption is made that every page read will be identified.
 - Since that assumption is made, different techniques are used for identifying documents.
- The document structure in the Learning Template approach is determined one of two ways.
 - Assume when using DemoSingleTIFFs that every page is a new document unless there is a doc separator in front of it.
 - Simulates documents that would come in at scan time
 - When documents are coming in from email or fax, there are no separator pages.
 - If there is a doc separator in front of it, it uses DemoMultiFormat to process variable sized documents that contain at least two pages.

More about the Learning Template workflow

- When images are pulled in, the Learning Template goes through what is called Managed Rotation
 - Therefore, it is important to turn off rotation when configuring image enhancement.
 - The Learning Template uses the Nuance document processing engine, which is powerful.
- As powerful as the Nuance engine is, if you have lots of images, the engine eventually crashes.
 - That situation is “normal” and expected
 - It recovers and continues processing
 - Nuance Instructions warn against running batches in an unmanaged fashion.
- At a minimum, you should always do deskew.

More about the Learning Template

- ‘Recognize Pages and Fields’ ruleset does full page recognition.
 - It should never need to be changed.
- ‘Create Documents’ is another U/I ruleset that you never need to change.
- Typically, in this type of application, only the main page is fingerprinted.
 - There are certainly exceptions, so if the characteristics of your images require that multiple pages be fingerprinted, by all means do so!
 - Use separator sheets between multi-page documents if your images require that.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-23. More about the Learning Template

What about routing and export?

- Routing is set up in the Built-in application.
 - There are no branches or child jobs.
- Do not ever do a quick export, which bypasses the verify step.
 - Always verify every document.
 - One of the reasons is that two clients might be confused because both are using Quick Books and their forms are almost identical.
 - With Learning Template applications it is a good practice to show every Main_Page to a user as part of verification.

Figure 6-24. What about routing and export?

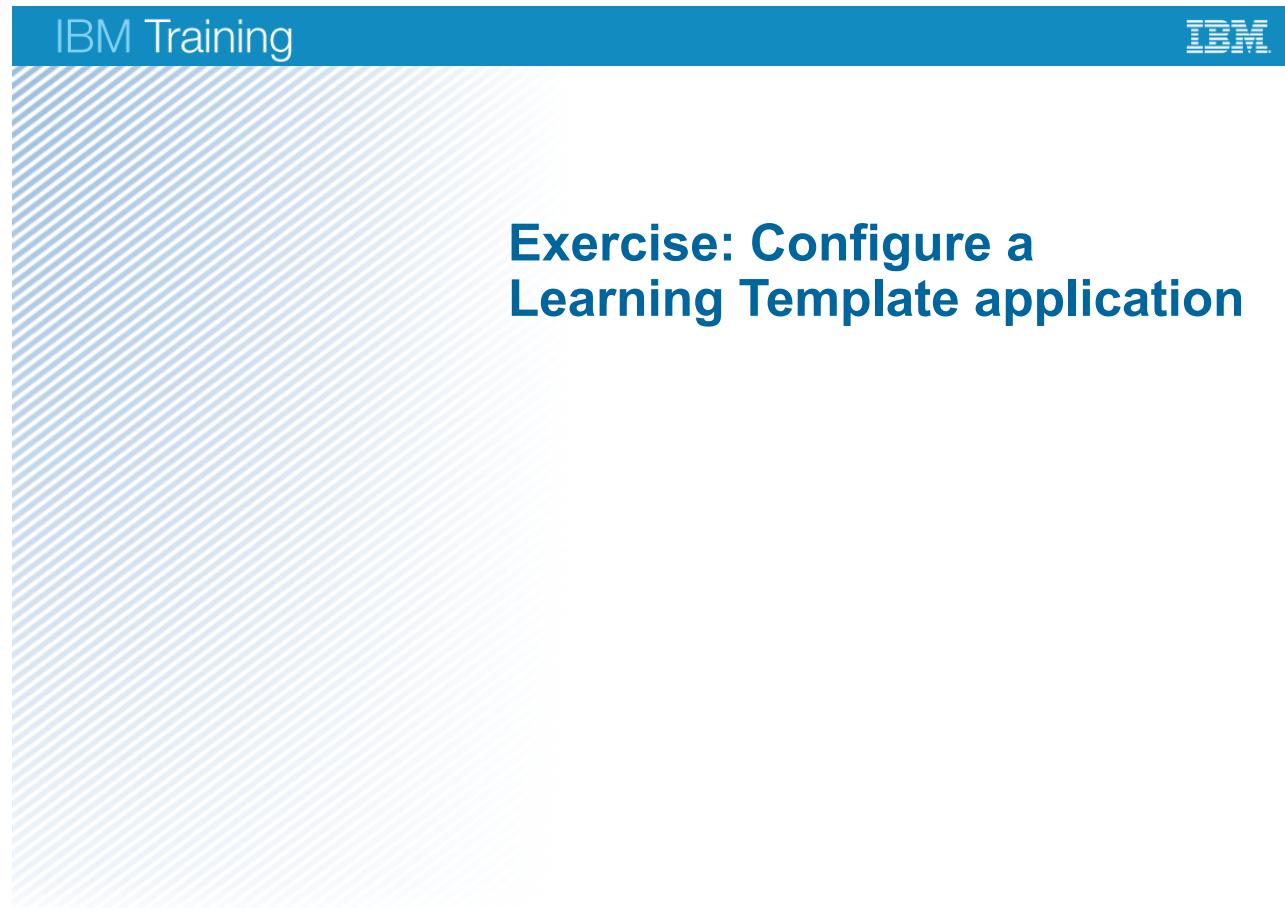
More things to know

- The Locate and Verify rulesets are built in to the Learning Template.
- You always have to define locate and validate rules:
 - It is a good practice to always validate every field you're capturing
 - At a minimum, specify:
 - If the field is required.
 - If it can be blank.
 - Whether it is overridable.
 - These options are easy to set in FastDoc.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-25. More things to know



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-26. Exercise: Configure a Learning Template application

Exercise objectives

- Basic application configuration

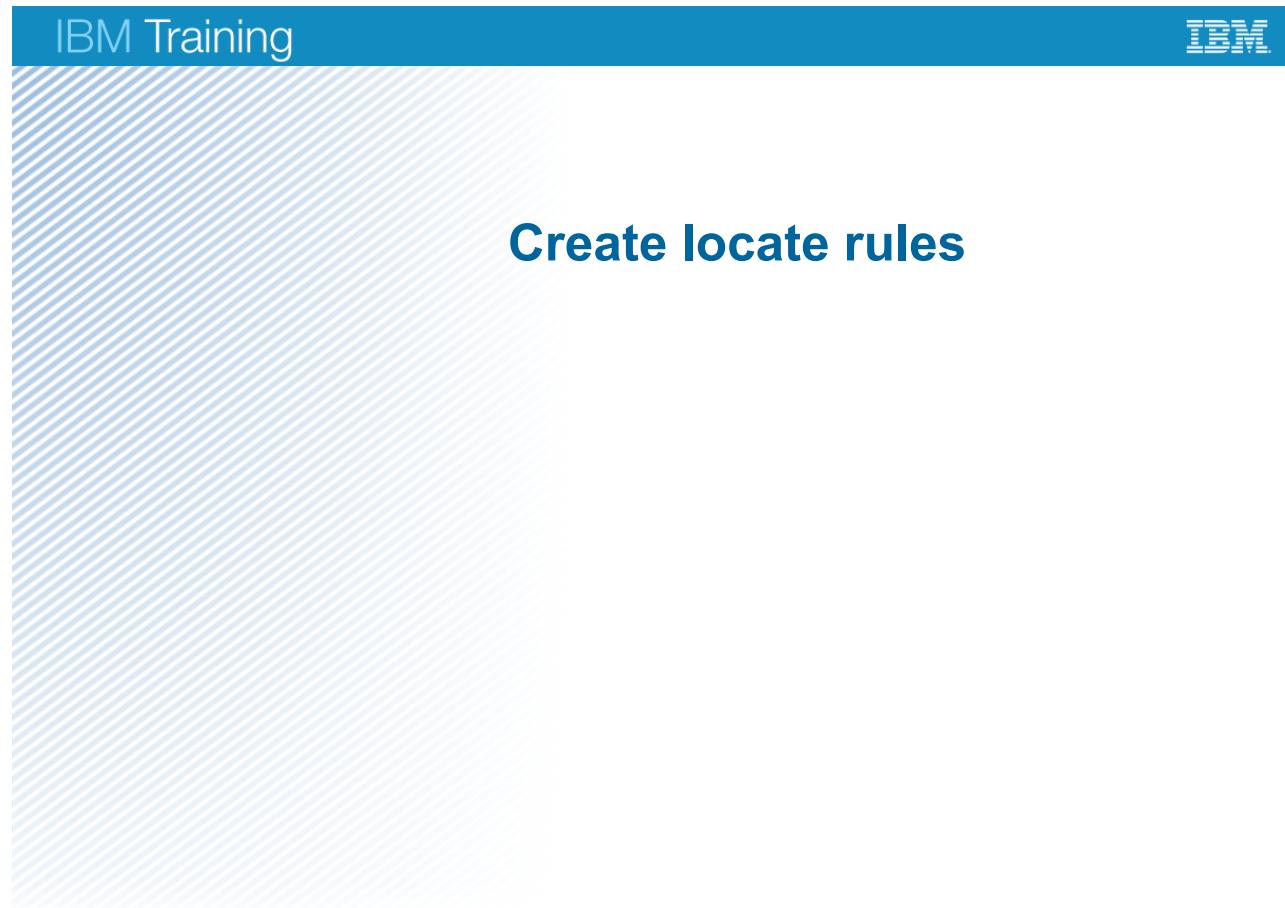


Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-27. Exercise objectives

Lesson 6.3. Create locate rules



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-28. Create locate rules

Lessons

- Create a Learning Template application
- Configure a Learning Template application
- Create locate rules
 - Create validate rules
 - Expand the locate rules
 - Expand the validate rules
 - Export to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-29. Lessons

Why is this lesson important to you?

- As a Datacap Solution/Application Builder, you build and deploy applications with the Datacap capture system and communicate solution details to the solution architect, administrator, and business analysts.
- When creating applications to process images that have inconsistent page layouts, you can save users time. You save them time by using locate rules to populate fields by using the click'n'key method when data positions move from one page to the next.

Figure 6-30. Why is this lesson important to you?

The basic locate actions

- Locate ruleset
 - Use key-word-search to locate and extract data from a document.
 - Use zones that are marked during fingerprinting to directly locate and extract values from cco files created by full page zonal recognition.
- Actions to read and extract data from zones:
 - ReadZones: Loads the position information for the current object and its children from the document hierarchy.
 - PopulateZNField: Populates the current field value in the page data file with the recognition data from the CCO file that lies within the field zone boundaries.
- If fields never changed locations, then these two actions would correctly read and extract data from the fields on a page.
- These actions are included in the built-in Locate rule for the Learning Template.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-31. The basic locate actions

Overview

A ruleset defines one or more rules that you can run on specific documents, pages, fields, or on the entire batch. You must associate a specific rule with one or more of the objects in the document hierarchy.

Key locate actions

The Locate ruleset is used to read and extract data from a document. Data can be directly read and extracted from the zones that are marked during fingerprinting the page or can be searched for using configured action and parameters.

Actions to read and extract data from zones:

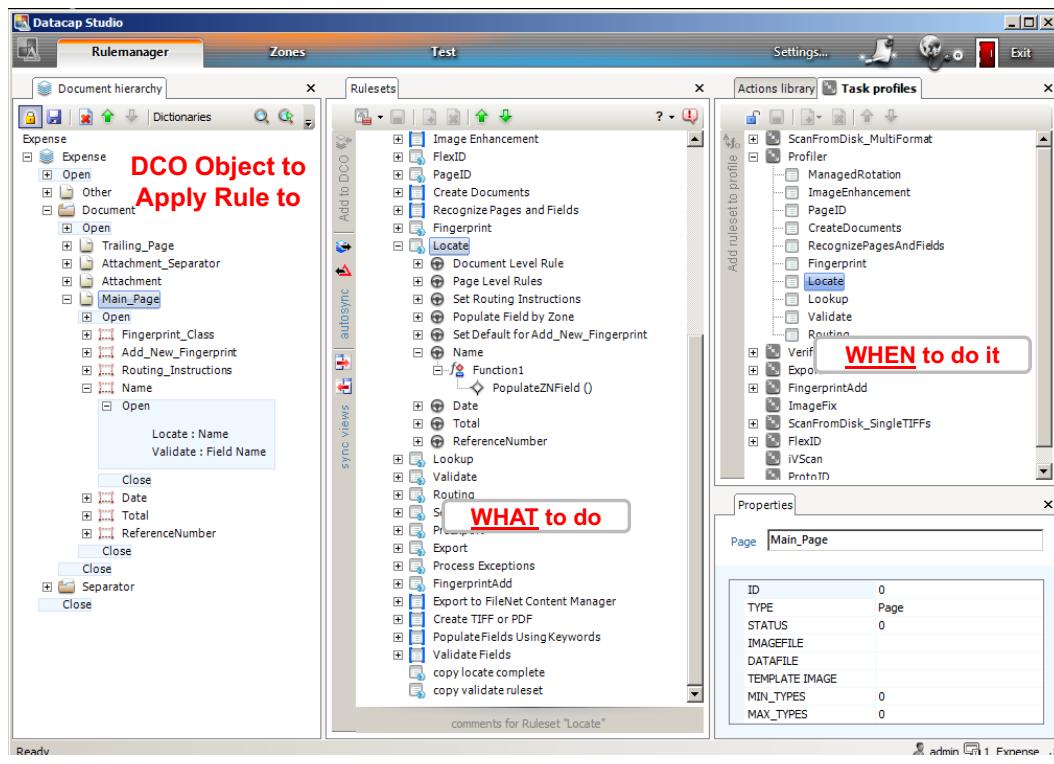
ReadZones: Loads the position information for the current object and its child objects from the document hierarchy.

PopulateZNField: Populates the current field value in the page data file with the recognition data from the cco file that lies within the field zone boundaries.

Advanced Locate actions are presented later in this course.

IBM Training

Rule configuration and run explanation



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-32. Rule configuration and run explanation

The image on the slide illustrates that:

- The Ruleset pane determines what must be done.
- The Task profiles determine when to do it.
- The document hierarchy pane determines which DCO objects to apply it to.

Locate actions (1)

Locate Actions	Description
FindKeyList	Locates the first (or next) occurrence of a word or phrase that matches one of the entries in a keyword file.
GoBelowWord	Moves down the specified number of lines from the previously found word or phrase.
GoLeftWord	Moves the specified number of words to the left of the previously found word or phrase.
GoRightWord	Moves the specified number of words to the right of the previously found word or phrase.
GroupWordsRIGHT	Groups words to the right of the previously found word if they are no more than the specified number of character widths apart.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-33. Locate actions (1)

Locate rule considerations

Locate rules can be used in place of fingerprinting to locate and extract image data. In this lesson, you develop locate rules for each image field and add them to the Locate ruleset. The Locate ruleset allows all of the fields to be identified on the image and for the data to be read. You must also associate the locate rules with the field definition in the DCO and define the Locate ruleset in the Task profile.

Locate Overview.

To recover data with locate rules, you need to create a rule for each field that is defined in the page definition on the Document hierarchy pane. Select fields that are not automatically populated with the two genetic rules that are already added to the Locate ruleset. Some of the advanced rules that can be used for this purpose are reviewed here. You are going to use the following Locate and Zone actions to find data.

Locate actions (2)

Locate Actions	Description
IsCurrency	Determines whether the value of the located word is a currency value (is numeric and includes a two-digit decimal amount).
IsDateValue	Determines whether the value of the located word is in one of the supported date formats.
MinLength	Determines whether the number of characters in the located word is less than or equal to the number specified (returns True if it is).
MaxLength	Determines whether the number of characters in the located word is greater than or equal to the number specified (returns True if it is).
RegExFind	Same as WordFind, but it supports regular expressions.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-34. Locate actions (2)

Locate and zone actions

Locate Actions	Description
UpdateField	Updates the current field in the page data file with the value and position of the located word.
WordFind	Locates the first (or next) occurrence of the specified word or phrase on the current page.

Zone Actions	Description
PopulateZNField	Populates the current field value in the page data file with the recognition data from the cco file that lies within the field zone boundaries.

Figure 6-35. Locate and zone actions



Rental_Agreement page

- Four fields to recognize and extract.

Thank you for renting with
Rental E-Receipt
Rental Agreement Number:703142974

Customer Information:	
Customer Name:	MILLARD BRYAN
Phone Number:	**158
Auto Purchase Discount Number:	LH4987
Customer Status:	Rental Car co. FIRST
Method of Payment:	AMEX
Credit Card #:	0000XXXXXX0000000000
Freq Thruster:	DL0999048766

Vehicle Information	
Car Group Rented:	FULL SIZE 4 DOOR
Car Type Chosen:	FULL SIZE 4 DOOR
Car Make Model:	AVN TOYO CHAIR 4DR
Plate Number:	MNHRU-0
Car Number:	Q287M3
Mileage Out: 5,078	
Mileage In: 5,083	
Mileage Overage: 174	
Fuel Out: 59	
Fuel In: 60	

Rental Information	
Pickup Date/Time:	18APR10/2115
Pickup Location:	MINNEAPOLIS ST. PAUL INTL APG 4400 40th Avenue NW ST PAUL, MN 55111 US 651/756-6220
Return Date/Time:	18APR10/2115
Return Location:	MINNEAPOLIS ST. PAUL INTL APG 4400 40th Avenue NW ST PAUL, MN 55111 US 651/756-6220

Vehicle Charges	
Time:	0 miles @ .20
	0.00
	TAX & 21% SURCHARGE
	26.20
	0 days @ 29.83
	0.00
	Fuel Service
	0.00
Total:	119.00
	Total Charge
	380.29
Time & Mileage	119.00
24120Y DPF	2.28
**10.00% FEE	18.01
LOW	0.00
FTR 0.00/DAY	3.13
3.325/DAY DPC	10.33
SUBTOTAL:	219.89
	Amount Due
	260.05

PROVISIONAL RECOVERY FEE
10.2% VEH TAX & 10% REPAIR FEE
EMERGENCY RECOVERY FEE - 100%
SUSPENSE FEE - 100%
FF 0.50/MTG (Excl 0.00)

Car Rental E-Receipt:
Please do not reply to this message.
If you have questions regarding this rental or if you do not wish to receive automatic receipts, please email e-receipts@ibm.com for assistance. If you have a question regarding the bill, please call us at 1-800-351-7988.
This receipt reflects your charges at the time of your return.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-36. Rental_Agreement page

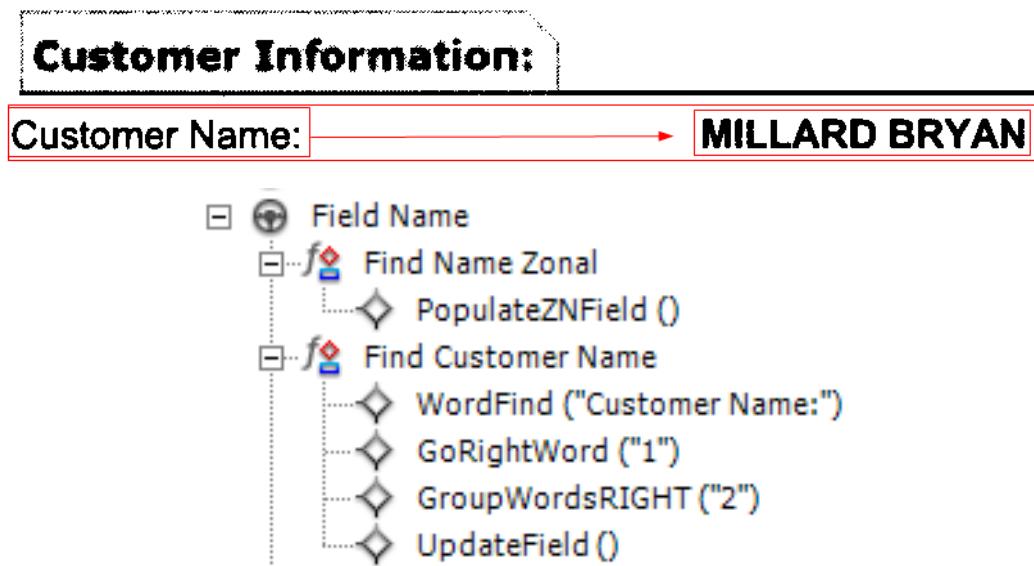
The screen captures show:

- A view of the Rental Agreement page with the four fields marked that you create locate rules for in the lesson lab.
- An expanded view of each of the four areas so they are more readable.
 - Rental Agreement Number
 - Customer Name
 - PickupDate/Time
 - Amount Due



Locate rule for customer name field

- Add Locate rules for Customer Name field to extract data.
- Find Name Zonal
- Find Customer Name



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-37. Locate rule for customer name field

The screen captures show:

- The section of the rental agreement with the customer name.
- The Field Name rule for locating the field on the page.

1. Find Name Zonal

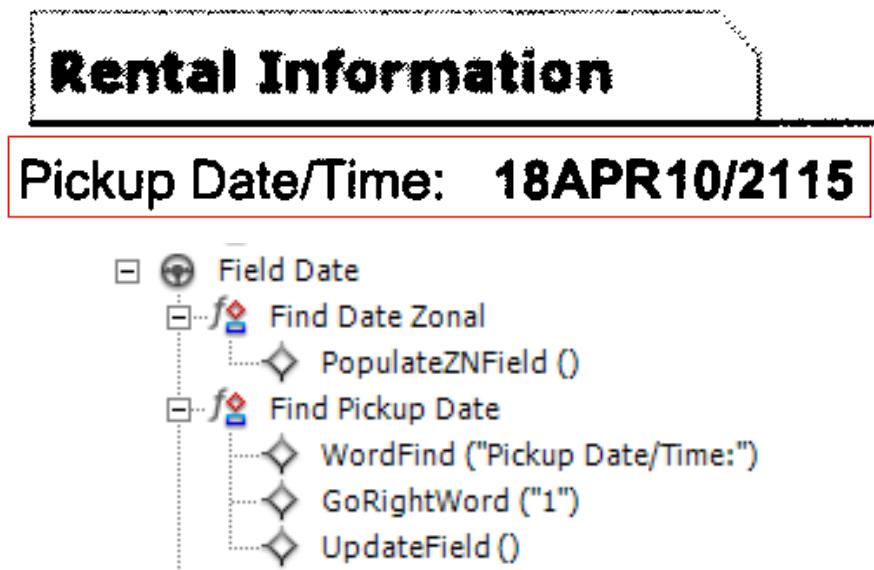
- Read the values from the zone (cco) file that is created by the full page zonal recognition action that is run at the page level.

2. Find Customer Name

- Search for the word "Customer Name:".
- Move right one word.
- Read two words to the right, which are the passenger first and surnames.
- The action updates the current field in the page data file with the value read.

Locate rule for pickup date

- Add Locate rules to individual fields to extract data.
- Find Reference Number Zonal
- Find Rental Agreement



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-38. Locate rule for pickup date

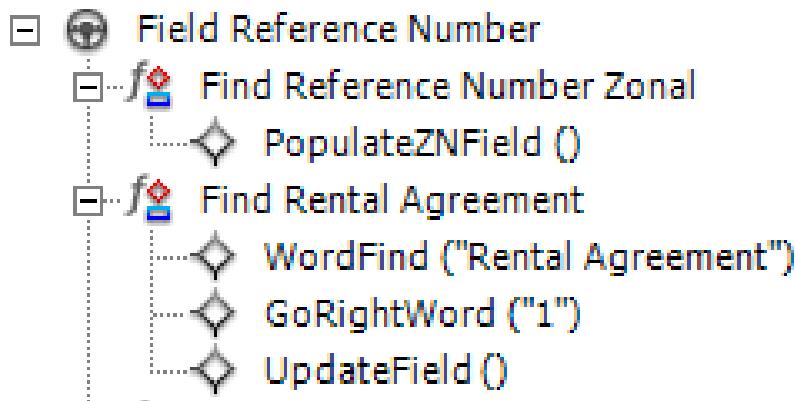
The screen captures show:

- The section of the receipt with the pickup date and time .
 - The Field Pickup Date rule for locating the field on the image.
1. Find Date Zonal
 - a. Read the values from the zone (cco) file that is created by the full page zonal recognition action that is run at the page level.
 2. Find Rental Agreement
 - a. Search for the text “Pickup Date/Time:”
 - b. Read one word to the right, which is the whole Expiration date field.
 - c. The action updates the current field in the page data file with the value read.
 - When this rule is complete you, do not have a valid date yet because the time is still appended. In the following lesson when you create the validate rules, you implement actions that correct this problem.

Locate rule for agreement number field

- Add Locate rules for Agreement Number field to extract data.
- Find Reference Number Zonal
- Find Rental Agreement

Rental Agreement Number:703142974



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-39. Locate rule for agreement number field

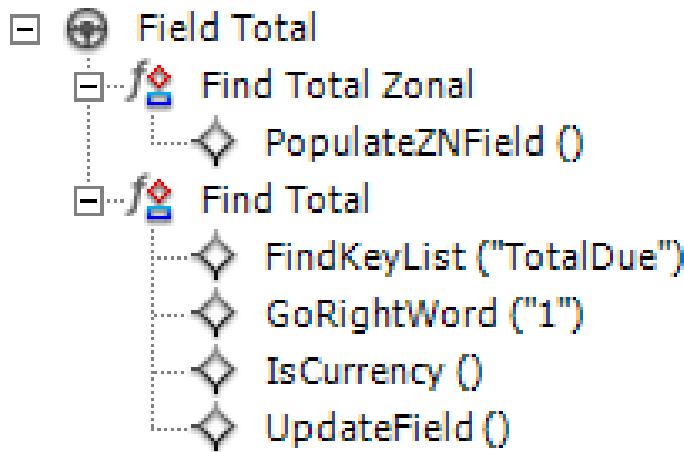
The screen captures show:

- The section of the page with the Rental Agreement Number.
 - The Field Reference Number ruleset for locating the field on the image.
1. Find Rental Agreement Zonal
 - a. Read the values from the zone (cco) file that is created by the full page zonal recognition action that is run at the page level.
 2. Find Rental Agreement
 - a. Search for the text “Rental Agreement”.
 - b. Read one word to the right, which is Number:*9digit number*.
 - c. The action updates the current field in the page data file with the value read.



Locate rule for total field

- Add Locate rules for total field to extract data.
- Find Total Zonal
- Find Total
- **Amount Due** 260.05



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-40. Locate rule for total field

The screen captures show:

- The section of the receipt with the Total amount.
 - The Field Total rule for locating the field on the image.
1. Find Total Zonal
 - a. Read the values from the zone (cco) file that is created by the full page zonal recognition action that is run at the page level.
 2. Find Total
 - a. Search for the words on the page that are a match any string that listed in the “TotalDue” key file. If a match occurs, move to the next action.
 - b. Read one word to the right, which is the amount due for the rental agreement.
 - c. Check that the value is in currency form.
 - d. Update the current field in the page data file with the value and position.
 - The TotalDue.Key file is in the C:\Datacap\<application name>\dco_<application name> folder.
 - It contains a list of many strings that are possible search strings for a value that might represent the key word that matches the page total value.

- The sequence of processing is, for each string of text that is read from the page, search the “.Key” file for a match. If a match is found process the next action.

Exercise: Create locate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-41. Exercise: Create locate rules



Exercise objectives

- Locate fields by using keyword search



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-42. Exercise objectives

Lesson 6.4. Create validate rules



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-43. Create validate rules

Lessons

- Create a Learning Template application
- Configure a Learning Template application
- Create locate rules
- Create validate rules
 - Expand the locate rules
 - Expand the validate rules
 - Export to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-44. Lessons

Why is this lesson important to you?

- Because you are working with the Learning Template, it is important to run validation rules on all the fields of the page you are configuring when you make your first pass through the application
- To successfully capture documents, you must be an expert at programmatically validating the captured data.

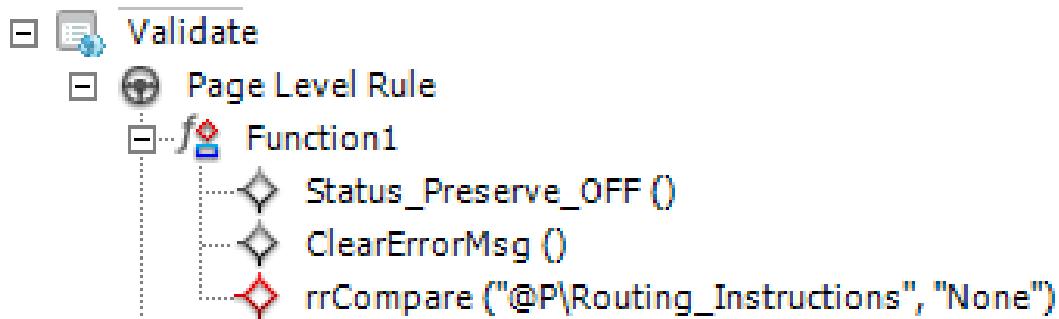
Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-45. Why is this lesson important to you?

The validate ruleset

- The Learning Template has a built-in page level rule that is applied to the Main_Page page.
- The Validate > Page Level Rule runs Status_Preserve_OFF
 - Setting Status_Preserve_OFF on the Validate Page rule allows the validation actions to assign a “problem” status to an invalid or empty field value.
 - Reported error allow verifiers to correct potential value errors



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-46. The validate ruleset

The screen capture shows: the ‘Page Level Rule’ uses the status_Preserve_OFF() action to ensure that field error are identified at the Verify step if errors occur.

Procedure to do simple validation and routing tasks.

Set Preserve OFF on Validate Page rule

In the default Validate Page rule, the action is set to Status_Preserve_OFF. You leave this rule unchanged. Having this rule present is important. It causes the batch to be routed to the Verify step if a field error is detected.

An Off condition of an object allows the actions of a Validate ruleset to assign a problem status to any field object with an invalid capture value. The Verify task Data Entry panel surrounds the value with a pink background, alerting the operator to the problem.

This error flagging occurs when a Validate field rule fails. The page status reflects the failure so that it shows at the Verification phase.

Add a field validation rule to a ruleset

Select the Validate ruleset in the Ruleset pane and lock the Ruleset structure for editing.

Right-click the validate Ruleset and click Add new rule. Change the rule name to identify the field that it operates on. For example, name the rule Field Name. The Field portion of the name identifies that the rule operates on a field node in the Document hierarchy. The Name portion of the rule name identifies that the field that is operated on by this rule is the name field.

Validate and export task relationship

- In addition to running validation as part of your first pass through the learning template application, it is important to also run export.
 - The location of each field on the Rental_Agreement page isn't saved until you run the export task.
- In the lab, you add a Field Validation Rule to each of the four fields in the application.
 - Some rules might require more than one function.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-47. Validate and export task relationship

Overview

It is critical to run export to update the location of the fields on your images. The field locations are not saved in a new fingerprint until export is complete.

Sample Validation actions

- IsFieldLengthMax (5)
 - Are there more than 5 characters in the field?
- IsFieldLengthMin (1)
 - Is there at least one character in the field?
- AllowOnlyChars (“0123456789”)
 - Allow only numeric digits 0-9 in the field
- Calculate (“SubTotal” + ‘Shipping’+ ‘Tax’ = ‘Total’)
 - Verify that the calculation based on field values is correct.
- IsFieldFilled ()
 - Is there any data in the field?
- IsFieldDate ()
 - Does the field contain a valid Date

Basic Learning Application

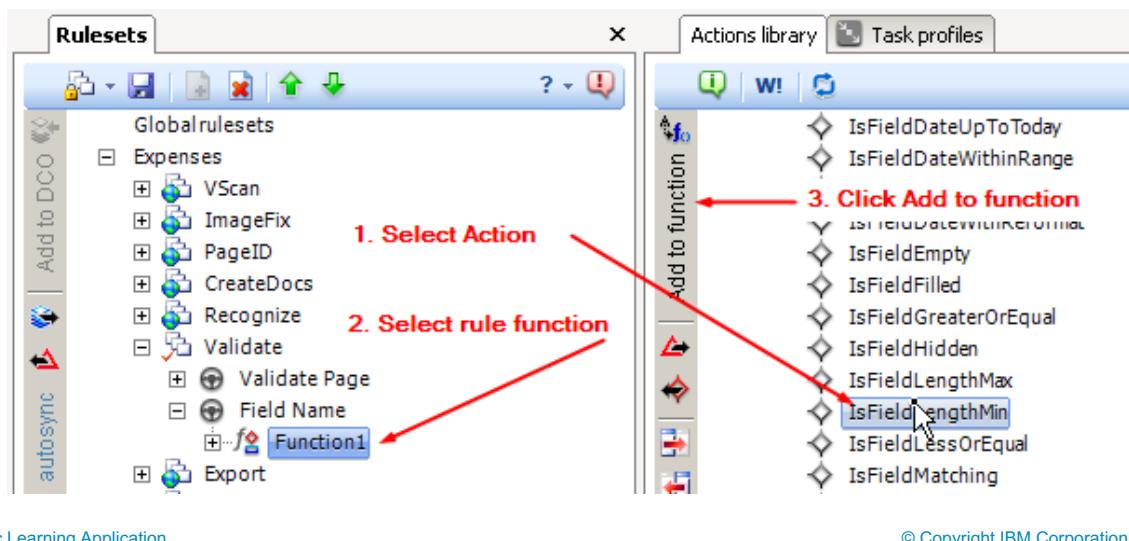
© Copyright IBM Corporation 2017

Figure 6-48. Sample Validation actions



Assign an action to the Field Name rule

- Select function
- Select action
- Add to function
- Set action properties



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-49. Assign an action to the Field Name rule

The screen capture shows a graphic representation off the method for assigning actions to functions in a rule.

Add an action that checks for a minimum field length of 5 in the name field.

Define a Validate rule for the Name field.

1. Select Function1 for the new Field Name rule that is selected in the Ruleset pane.
2. Select the IsFieldLengthMin action in the Validations Actions group of the Actions library pane.
3. Click the Add to function link.
4. Click the IsFieldLengthMin() action that you just added to the rule and set the Parameter to 5 on the properties pane.

Validations actions

Locate Actions	Description
AllowOnlyChars	Removes all characters that are not specified as allowed characters.
FailRuleSet	Causes the entire ruleset to fail.
IsFieldDate	Returns True if the current field value is a date.
IsFieldDateWithinRange	Returns True if the current field value is a date within the specified range.
IsFieldDateWithReformat	Returns True if the current field value is a date and reformats the value to the date format specified.
IsFieldFilled	Returns True if the current field value is not empty.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-50. Validations actions

Validate and route

Overview

When field data is located and extracted from the page image, it is necessary to validate the reliability of the data. This information is used to be sure that all fields are completed with an acceptable degree of accuracy. Validate rules are used at verify time to make sure that the operator does not enter invalid values.

Some of the actions that can be used for validation are reviewed here.

Validations actions

Validate Actions	Description
IsFieldLengthMax	Returns True if the length of the current field value is less than or equal to the length specified.
IsFieldLengthMin	Returns True if the length of the current field value is greater than or equal to the length specified.
LeftTruncate	Deletes characters from the end of the current field value until the length equals the length that is specified.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-51. Validations actions

Mixing locate and validate actions

- It is acceptable to mix Locate and Validation actions in the same function.
- Locate actions operate on cco fingerprint values.
- Validation actions act on the field value.
- When mixing actions, the UpdateField() action must follow the locate actions; then the Validation actions follow UpdateField().
 - If this sequence is not followed, then the validation actions act on an uninitialized value.
- Reason for not mixing Locate and Validation actions.
 - Locate rule runs only in the Profiler task.
 - Validate rule runs in the Profiler and Verify task.
 - Any validation actions included in the Locate rule are not run at the Verify step if you do run Validations.

Figure 6-52. Mixing locate and validate actions

Exercise: Create validate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-53. Exercise: Create validate rules

Exercise objectives

- Configure field validation rules



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-54. Exercise objectives

Lesson 6.5. Expand the locate rules

Expand the locate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-55. Expand the locate rules

Lessons

- Create a Learning Template application
 - Configure a Learning Template application
 - Create locate rules
 - Create validate rules
-  Expand the locate rules
- Expand the validate rules
 - Export to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-56. Lessons

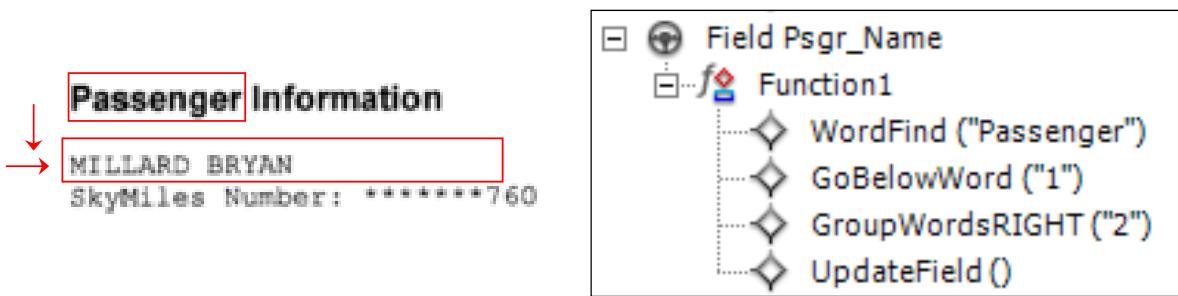
Why is this lesson important to you?

- As a Datacap Solution/Application Builder, you build and deploy applications with the Datacap capture system and communicate solution details to the solution architect, administrator, and business analysts.
- To build Datacap applications, you need to be an expert in the use of the locate actions. Locate actions are an alternative and sometimes more explicit method of identifying fields on a page.

Figure 6-57. Why is this lesson important to you?

Locate rule for passenger name field

- Add Locate rules for passenger name field to extract data.
 - Search for the word “Passenger”.
 - Step down one line.
 - Read two words to the right, which are the passenger first and last names.
 - The action updates the current field in the page data file with the value.



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-58. Locate rule for passenger name field

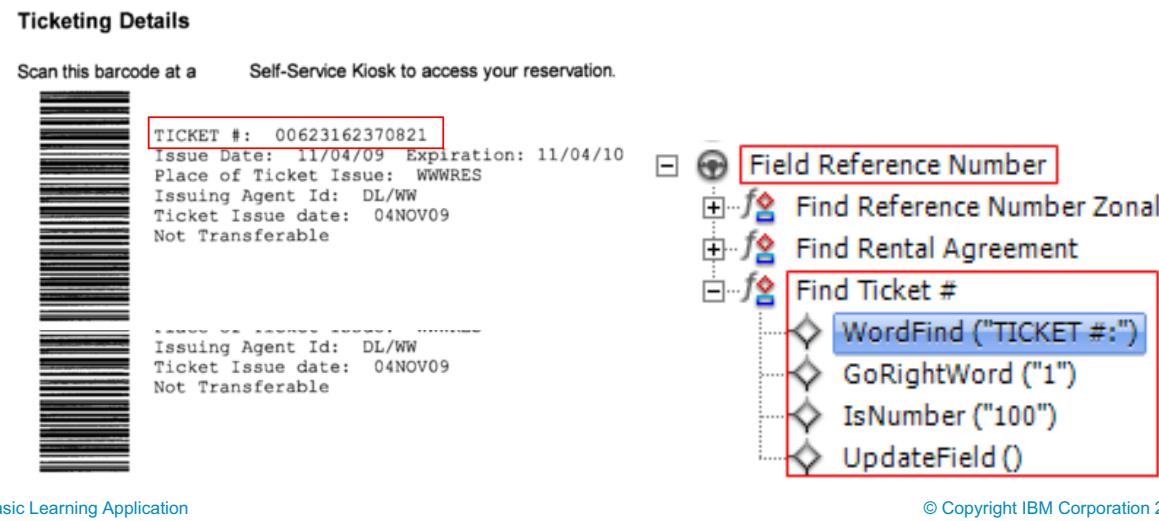
The screen captures show:

- The section of the receipt with the passenger name.
- The Field Psng_Name rule for locating the field on the image.



Locate rule for the ticket number field

- Add Locate rules for ticket number field to extract data.
- Find Ticket #
 - Search for the text “TICKET #”.
 - Read one word to the right,



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-59. Locate rule for the ticket number field

The screen captures show:

- The section of the receipt with the ticket number.
- The Field Ticket Number rule for locating the field on the image.

Find Ticket

- Search for the text “TICKET #”.
- Read one word to the right, which is the 14-digit ticket number.
- The action updates the current field in the page data file with the value.



Locate rule for issue dates

- Add Locate rules to individual fields to extract data.
 - Search for the text “Issue Date:”
 - Read one word to the right, which is the Issue Date field.
 - Verify that the value is a date.
 - The action updates the current field in the page data file with the value.

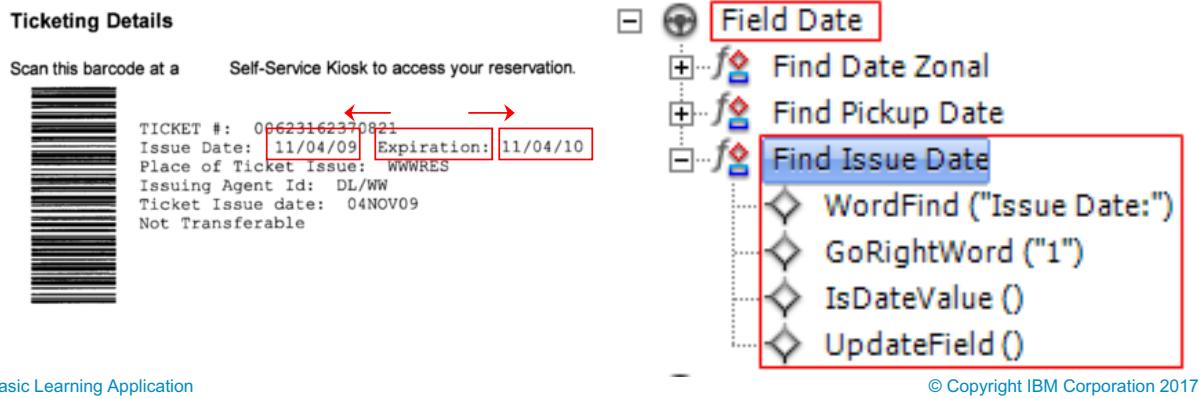


Figure 6-60. Locate rule for issue dates

The screen captures show:

- The section of the receipt with the Issue date and the Expiration date.
- The Field Issue Date rule for locating the field on the image.

Exercise: Expand the locate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-61. Exercise: Expand the locate rules

Exercise objectives

- Expand the Locate Ruleset for Air Receipt pages
- Create an Air_Receipt page fingerprint

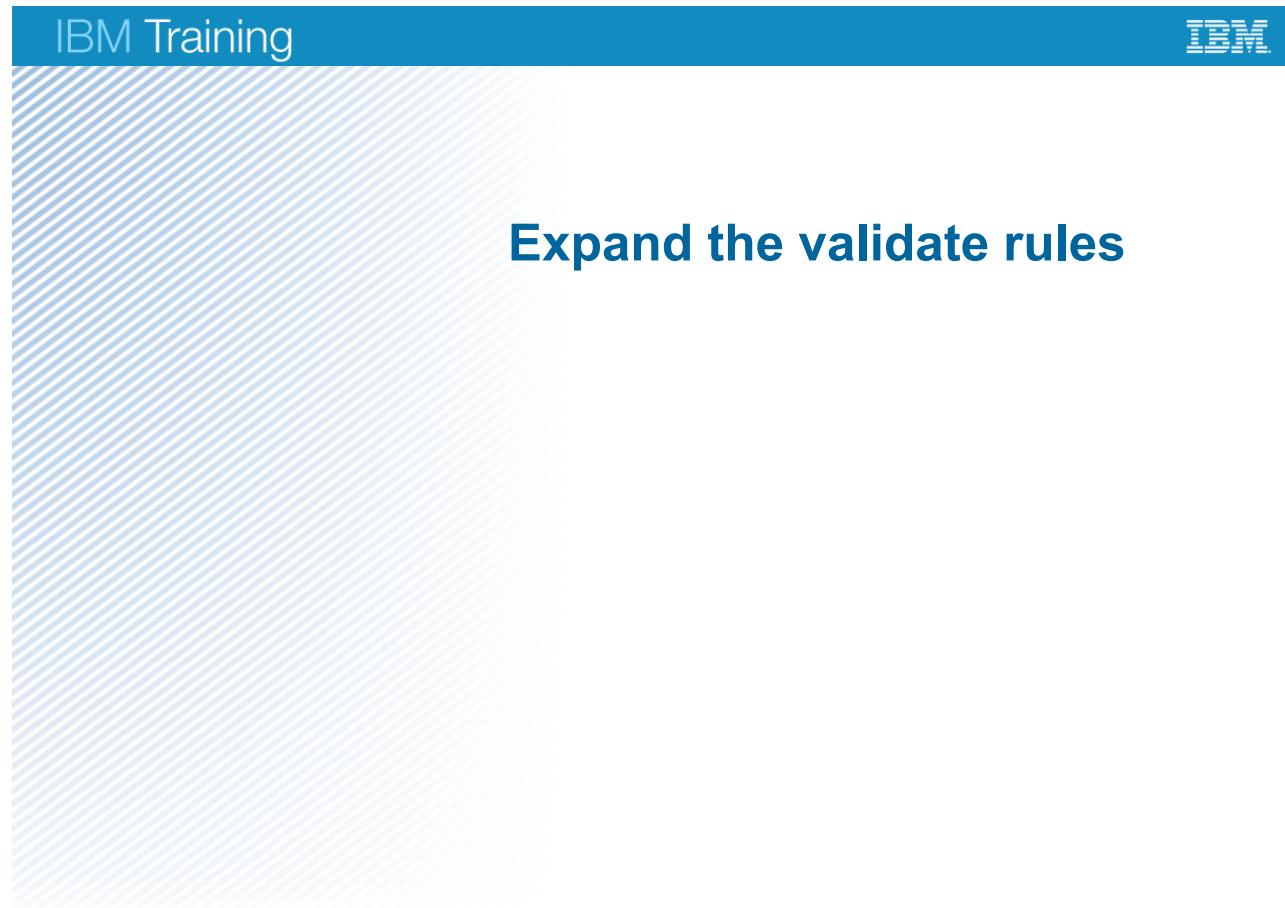


Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-62. Exercise objectives

Lesson 6.6. Expand the validate rules



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-63. Expand the validate rules

Lessons

- Create a Learning Template application
 - Configure a Learning Template application
 - Create locate rules
 - Create validate rules
 - Expand the locate rules
 - Export to an IBM FileNet repository
-  Expand the validate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-64. Lessons

Why is this lesson important to you?

- As a Datacap Solution/Application Builder, you build and deploy applications with the Datacap capture system and communicate solution details to the solution architect, administrator, and business analysts.
- To build an efficient Datacap application That saves users time, you need to be expert in the use of the validate actions because they are key to identifying data errors on captured documents.

Figure 6-65. Why is this lesson important to you?



Validate rule for the issue dates

- Add a rule to validate field data for both dates.
 - Is the field exactly 8 characters long?
 - Is the value a valid date in one of the standard date formats.

The screenshot displays a 'Ticketing Details' screen. On the left, there's a barcode with the text 'TICKET #: 00623162370821' above it. Below the barcode, the 'Issue Date' and 'Expiration' fields are highlighted with a red border. To the right of the barcode, the text 'Self-Service Kiosk to access your reservation.' is visible. Further down, ticket details are listed: 'Place of Ticket Issue: WWWRES', 'Issuing Agent Id: DL/WW', 'Ticket Issue date: 04NOV09', and 'Not Transferable'. On the right side of the screen, a hierarchical tree view shows validation rules. The root node is 'Field Date', which branches into 'Issue Date' and 'Pickup Date'. 'Issue Date' is selected and highlighted with a blue border. Under 'Issue Date', three actions are listed: 'IsFieldLengthMin ("8")', 'IsFieldLengthMax ("8")', and 'IsFieldDate ()'. A red box encloses the entire validation rule section.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-66. Validate rule for the issue dates

The screen captures show:

- Page image format for the Issue and Expiration date
- Validate Date Rule

Validate date

How do you tell a valid date on the two Air_ Receipt date fields?

If you are dealing with an eight-character format, 11/11/11, you can use the IsFieldLengthMax and IsFieldLengthMin actions. Then, you can use the IsFieldDate action to verify that the three components of the date are in range.

Another action that you can use to validate that a date is within a specific date range is the IsFieldDateWithinRange action. This check is not valid for the scenario that you are dealing with, so you are not going to use this action.

Depending on your requirements, you can control data tightly. For example, you can verify that the expiration date is not before the issue date.

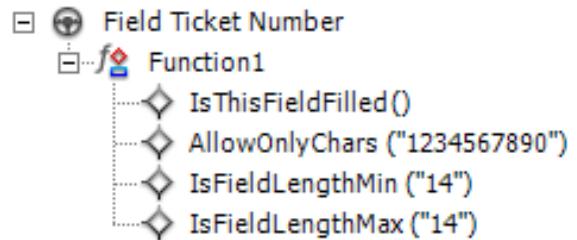
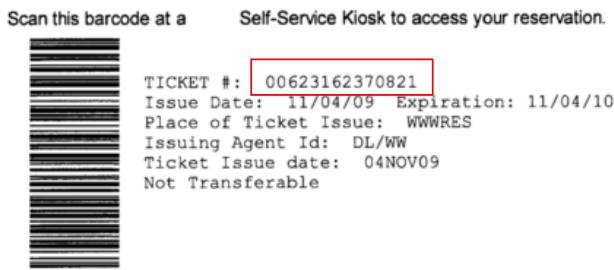
For this demonstration, keep it simple, using only IsFieldLengthMax, IsFieldLengthMin, and IsFieldDate. The same rule is used for both date fields.



Validate rule for the ticket number field

- Add a rule to validate field data for the ticket number.
 - Does field contain any data?
 - Is the data limited to decimal digits?
 - Is the field size exactly 14 digits?

Ticketing Details



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-67. Validate rule for the ticket number field

The screen captures show:

- Page image format for the Ticket number.
- Validate Ticket Number Rule

Verify

- Operator reviews pages to confirm and correct problem fields.
 - If validation errors were detected with Routing rules.
 - If verification is forced by using the Routing Force Verify rule.
- Documents with no errors can bypass operator review.
 - If, on the page, the DCO status is not set with the Routing Force verify rule.
- You can fill empty fields by the click and key method.
 - If a full page OCR operation is done on the page.
 - To do click and key, click the field on the image next to the form.
- Correct erroneous data.
 - Manually verify image data and retype it on the form.
- Manually verify all fields on the form.
 - Uses Alt V to force verification after changes.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-68. Verify

Overview

During verification, Datacap shows pages to an operator to confirm, and if necessary, correct problem fields. You have control over whether documents with no validation errors are shown. Problem fields include:

- Character fields with one or more low-confidence characters
- OMR fields with low-confidence values
- Fields with validation errors

Exercise: Expand the validate rules

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-69. Exercise: Expand the validate rules

Exercise objectives

- Expand the validate rules for the Air_Receipt pages



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-70. Exercise objectives

Lesson 6.7. Export to an IBM FileNet repository



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-71. Export to an IBM FileNet repository

Lessons

- Create a Learning Template application
- Configure a Learning Template application
- Create locate rules
- Create validate rules
- Expand the locate rules
- Expand the validate rules

 Export to an IBM FileNet repository

Why is this lesson important to you?

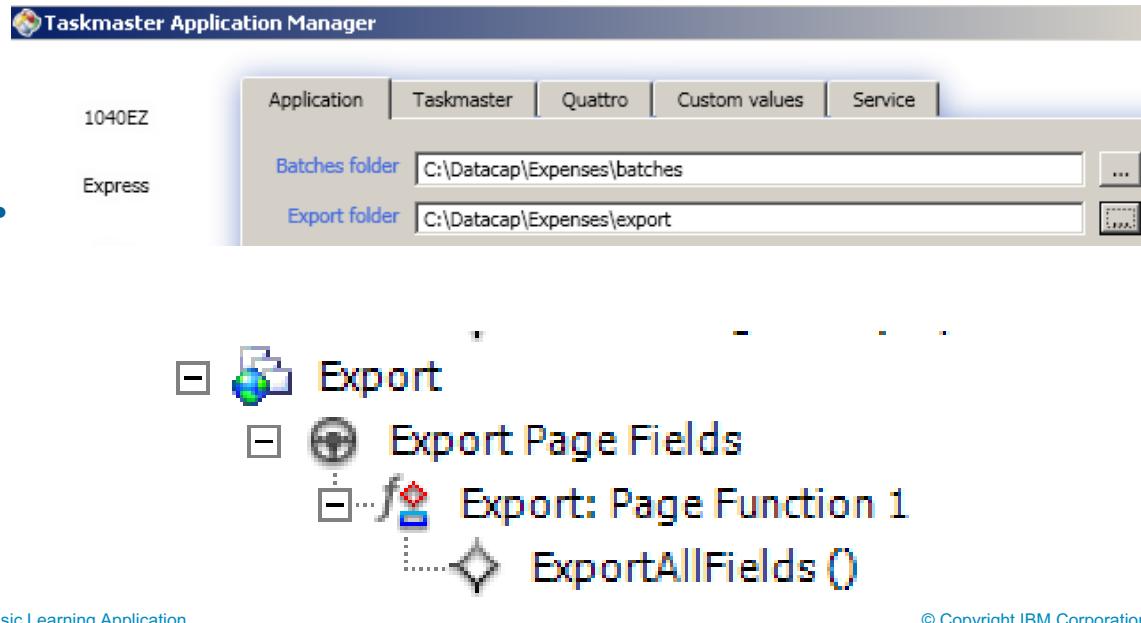
- As a Datacap Solution/Application Builder, you build and deploy applications with the Datacap capture system and communicate solution details to the solution architect, administrator, and business analysts.
- To build a complete Datacap application, you must be able to complete the cycle by exporting the final captured documents to the selected content store. In this lesson, you export to an IBM FileNet P8 repository.

Figure 6-73. Why is this lesson important to you?



Export types

- Export data to a text file, an XML file, a database, a document management system, or a custom business process.
- Default Export Folder - Exports folder for the application.



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-74. Export types

The images on the slide show:

- The Export folder parameter in the Datacap Application Manager.
- The Export ruleset that is used to export all images to text file.

Overview

The Export step is the last step of the capture process. This step determines where the final capture data is placed and what form it is in when it is exported.

Export to a text file

Datacap can export data to a text file, an XML file, a database, a document management system, or a custom business process. The default output format is a text file. The text file is the form that you are going to use in this section. Other export data forms are introduced in later sections of this class. The default location to place the output is in the exports folder for the application. Another location, the FileNet P8 repository, is introduced in a later lesson.

Default Export Folder

The default export folder is defined in the Datacap Application Manager by the Export folder parameter.

The default Export Page Fields rule stipulates that all identified fields are exported.

Preparation for export to IBM FileNet repository

- Create a folder structure to receive diverse document types if that is what you are exporting.
- Define a document class for each document type.
 - The document class requires a property that is defined for each field that is extracted from the captured document.
 - If the same field data types are extracted from all of the image types, then the same document class can be used to store the documents on the IBM FileNet repository.
 - The Expenses DocClass is for both the Rental_Agreement and Air_receipt in the lab.
 - You might add a property that indicates document type, that is, Rental_Agreement or Air_Receipt but the labs don't do that.
 - It is more likely that different document classes are preferred for each captured document type.

Figure 6-75. Preparation for export to IBM FileNet repository

Overview

Datacap applications can export captured data and images to a flat text file or to other external repositories with standard actions and separately licensed connectors. The external repositories include: systems and databases, IBM repositories, and other vendor repositories. In an earlier lesson, data was exported to a text file. In this exercise, you set up rules to export to an IBM FileNet repository.

Some preparation must be done for exporting to an IBM FileNet repository. Define a document class for a set of page data that is written to the object store. The document class must have a property that is defined for each value that is extracted from the page. Define three export rules. One rule is a batch rule that establishes a connection to the object store that the export data is written to. The other IBM FileNet export rules are one for each of the document sets that the Expense application processes.

Actions to export to IBM FileNet repository

FileNetP8 Action	Description
FNP8_Login	Sets the user ID and password for login to the FileNet P8 system.
FNP8_SetDestinationFolder	Sets the destination folder for the documents that are uploaded.
FNP8_SetTargetClassID	Sets the location type where docs will be stored, that is, ObjectStore
FNP8_SetTargetObjectID	Sets the FileNet P8 object store for the uploaded files.
FNP8_SetLocale	Sets the locale, that is, en_US
FNP8_SetURL	Sets the URL for the FileNet P8 web service.

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-76. Actions to export to IBM FileNet repository

The actions for exporting data to a FileNet P8 repository are in the FileNet P8 group.



FastDoc batch level configuration

- Learning and Form Template both use the Export to FileNet Content Manager ruleset.

Ruleset **Export to FileNet Content Manager**

FileNet Content Manager Export Settings

Batch Information

FileNet Content Manager URL:*	<input type="text" value="http://ecmedu01:9080/wsi/FNCEWS40MTOM/"/>
User ID:*	<input type="text" value="p8admin"/>
Password:*	<input type="password" value="*****"/>
Confirm password:*	<input type="password" value="*****"/>
Locale:	<input type="text" value="English (United States)"/>
Storage object id:*	<input type="text" value="DCExport"/>
Parent folder:	<input type="text" value="/Expenses/"/>
Subfolder to create for batch:	<input type="text" value=""/>
Number of upload attempts:	<input type="text" value="0"/>
Upload timeout:	<input type="text" value="600000"/>

Basic Learning Application © Copyright IBM Corporation 2017

Figure 6-77. FastDoc batch level configuration

The screen captures show:

- The FastDoc interface for batch level configuration for the Export to FileNet Content Manager ruleset.
- The Datacap Studio definition representation of the document level configuration.

FileNet Content Management URL

- The URL for remote access to the IBM FileNet Content Manager Repository

User ID

- The user id used to log in the IBM FileNet Content Manager server.

Password and Confirm Password

- The password used to log in the IBM FileNet Content Manager server.

Storage object id

- The repository name

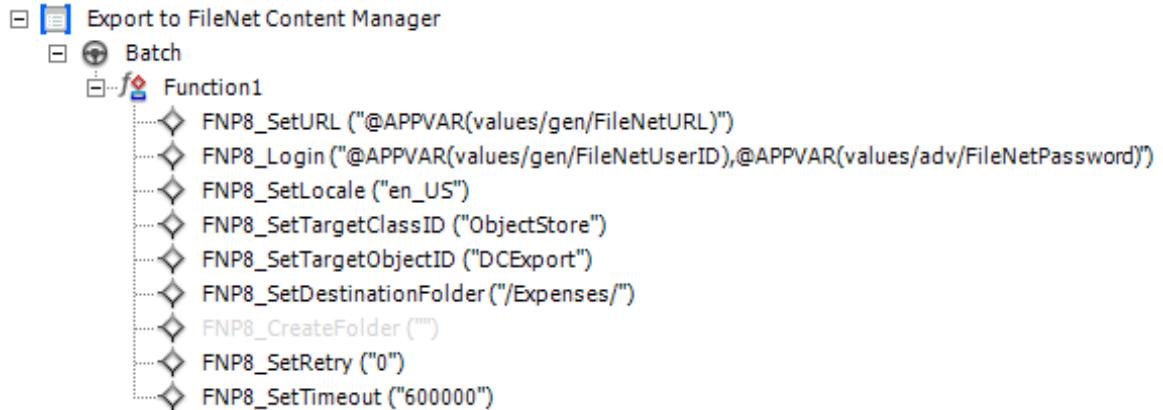
Parent folder

- The folder where Datacap documents are exported to.



Function to connect to the IBM FileNet server

- Learning and Form Template default batch level rule.
 - It shows configuration for the student image lab exercise.



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-78. Function to connect to the IBM FileNet server

The screen captures show:

- The Datacap Studio representation of the document level configuration.

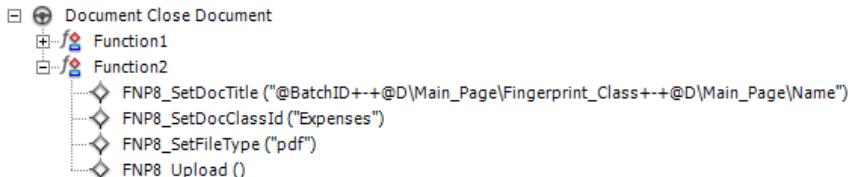


FastDoc document level configuration

- Learning and Form Template document level configuration the Export to FileNet Manager ruleset.

The screenshot shows the 'FileNet Content Manager Export Settings' configuration screen. At the top, there is a dropdown menu labeled 'Ruleset' set to 'Export to FileNet Content Manager'. Below the dropdown are three input fields: 'Document title:' containing '@BatchID++@D\Main_Page\Fingerprint_Class++@D\Main_Page\Name', 'Document class ID:' containing 'Expenses', and 'Document file extension:' containing 'pdf'. A section titled 'Document Information' is expanded, showing the text 'The Document document will be exported.'

- Datacap Studio view of document level configuration



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-79. FastDoc document level configuration

The screen captures show:

- The FastDoc interface for document level configuration.
- The Datacap Studio representation of the document level configuration.

Document title parameter analysis

- @BatchID - The same as the folder name in the batches folder. Example 20160829-000004.01
- @D\Main_Page\<field name> - this example provides access any field in the Main_Page page.
- Fingerprint_Class examples for this class are Rental_Agreement and AIR_Receipt.

Document Class ID

- Is the document class as defined in the IBM FileNet repository.

Document file extension

- The combination of the Batch ID and the Document file extension.
- The document must already be conversion to the file type that you specify in this parameter.
- In the lab exercise for this lesson, you use the CreateTIFF or PDF ruleset to produce a pdf to document to export.



FastDoc field level configuration

- For each document property to write to IBM FileNet repository backend server, you must configure a field level rule for the Export to FileNet Content Manager ruleset.

The screenshot shows the 'FileNet Content Manager Export Settings' configuration screen. At the top, there is a dropdown menu labeled 'Ruleset' with 'Export to FileNet Content Manager' selected. Below the dropdown are three buttons: a blue square, a green folder icon, and a white square with a minus sign.

The main section is titled 'FileNet Content Manager Export Settings'. It contains a collapsed section 'Field Information' indicated by a triangle icon and a checked checkbox. The sub-section details are as follows:

- Symbolic name:** Employee_Name
- Property type:** String
- Multiple value property:**

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-80. FastDoc field level configuration

The screen captures show:

- The FastDoc interface for field level configuration.

Symbolic name:

- The name of the document property that is defined in the IBM FileNet repository Document Class.

Property Type:

- This type that is defined for the Datacap field type must match the type of the property that is defined in the IBM FileNet repository Document class.

Setting the value of a document's properties

- Function to set field property values
- You must specify the field name on the IBM FileNet repository server and the data type
- In addition, you must pass in the field value



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-81. Setting the value of a document's properties

The screen captures show:

- The Datacap Studio representation of the field level configuration.

Definition for the **FNP8_SetPropertyEx** (StrParam) action.

Parameters: comma-separated string values:

1. Property ID: The name of an existing document property in the FileNet library (equivalent to a document index field)
2. String value or a predefined Smart Parameter variable to assign to the property. @F represents the field that this rule is associated with.
3. Optional property type: If this parameter is not specified, the property type defaults to a 'string'. Supported types are Binary, Boolean, DateTime, Float, ID, Integer, Object, and String.

Configure multi-page documents

- Learning Template assumes one page per document.
- To group multiple pages into a document use separators page.
- General rule:
 - Pages are treated as one page documents until a separator page it encountered.
 - Then all pages are assumed to be in the same document until another separator page is encountered.
- Sample multi page batch:
 - Separator page.tif
 - Car1.tif
 - Air1 sample1.tif
 - Hotel1.tif
 - Separator page.tif

Figure 6-82. Configure multi-page documents

The screenshot shows a blue header bar with the text "IBM Training" on the left and the "IBM" logo on the right. Below the header, the title "Separator page sample" is displayed in bold blue text. To the left of the main content area, there is a bulleted list:

- Adjust Image Enhancement
- Set the Remove Lines > Minimum

The main content area contains a document separator sheet. At the top, there are three thick horizontal black bars. Below them, the text "Document Separator Sheet" is printed in a standard font. Underneath the text is a barcode.

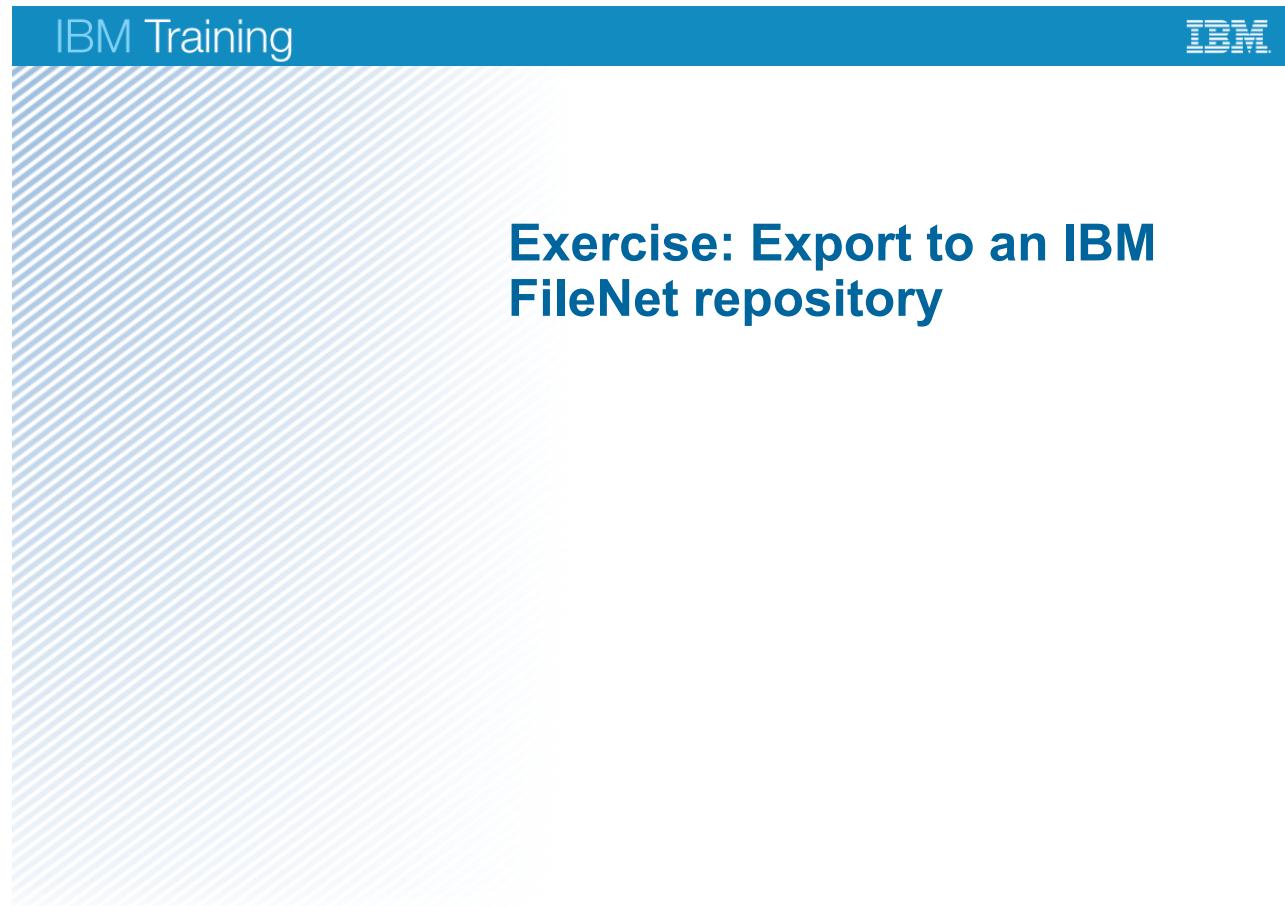
Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-83. Separator page sample

The screen capture shows a sample separator page with a horizontal bar, The words "Document Separator Sheet", and a vertical bar code.

- When you first configure your multi-page batch for a Learning Template application, you might detect the following behavior:
 - The separator pages are treated as standard page.
- The cause of this is probably because the Separator page bare code is being corrupted by the Image Enhancement ruleset.
- To correct this condition, set the Minimum length parameter under the Remove Lines option to a value ≥ 70 .
- Expense > Document > Main_Page > Image Enhancement > Remove Lines > Minimum length



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-84. Exercise: Export to an IBM FileNet repository

Exercise objectives

- Export Expense documents to an IBM FileNet repository



Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-85. Exercise objectives

Unit summary

- After completing this unit, you should be able to:
- Create a basic learning application with the Datacap Application wizard
- Apply image enhancement techniques and field recognition potential
- Define a document structure
- Create and expand Locate rules
- Create and expand Validate rules
- Export the scanned documents to an IBM FileNet repository

Basic Learning Application

© Copyright IBM Corporation 2017

Figure 6-86. Unit summary

Unit 7. Debug and Test

Estimated time

02:00

Overview

This unit introduces you to the tools and interfaces that are used for debugging and testing Datacap applications.

How you will check your progress

- Successfully complete the activities in the Student Exercises book.

References

IBM Knowledge Center

http://www.ibm.com/support/knowledgecenter/SSZRWV_9.0.1/com.ibm.datacaptoc.doc/datacap_9.0.1.htm

Unit Objectives

- After completing this unit, you should be able to:
- Configure and access logs for debugging
- Use the Datacap Studio Test tab to step through a batch for testing

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-1. Unit Objectives

Lesson 7.1. Application Debugging

Application Debugging

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-2. Application Debugging

Topics

- ▶ Application Debugging
 - Datacap Studio Test Tab

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-3. Topics

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users. As part of that process, you must be able to debug solutions.
- You can run an application from Datacap Studio and monitor it during execution, then determine whether the rules are running as you expect.
 - Become familiar with the debugging features that are included with the product by intentionally creating errors in your application.

Figure 7-4. Why is this lesson important to you?

Application Debugging and Datacap Studio

- Datacap Studio includes integrated debugging functionality.
 - Through this functionality you can control execution of your application.
- You can run an application from Datacap Studio.
 - You're able to monitor the application during execution.
 - You can also determine whether the rules are running as you expect.
- Application debugging requires that you review at least two runtime log files.
 - The Rulerunner Service (RRS) log.
 - The Task Log.
- There are also a few other logs that are useful depending on what the issue is with your application.

Figure 7-5. Application Debugging and Datacap Studio

Application debugging requires that you review two runtime log files, which are the Rulerunner Service (RRS) log and the task log. The RRS log provides detailed information about each action and is most helpful to application developers. The task log documents internal calls and is used mostly by IBM software support.

Datacap Studio includes integrated debugging functionality through which you can control the execution environment and monitor your application at runtime.

The Logs Used in Debugging

- As Rulerunner runs each action, it writes detailed logging information to a Rulerunner Service (RRS) log file.
- The Rulerunner Service Logs = RRS Logs:
 - They provide detailed information about each action as it is executed.
 - These logs are most helpful to application Developers.
 - There is a global RRS log for each execution of an application as well as individual RRS logs for components like TM Server, Rulerunner server, and TM Web.
- The Task Log:
 - Documents internal calls.
 - Is used mostly by IBM software support.

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-6. The Logs Used in Debugging

Datacap log files

Datacap generates two types of log files during task execution.

- Rulerunner Service (RRS) log files include detailed information about each action as it runs.
- Task log files document mostly internal calls and is most helpful to IBM Software support.

Additionally, Report Viewer and Rulerunner can generate their own log files.

The RRS Log Files

- Rulerunner also generates an RRS log file whenever you run a task from Datacap Studio.
 - RRS task log files are named `task_rrs.log`, ie. `verify_rrs.log`.
- You can set Rulerunner logging by application and task.
- Logging is enabled by default for all tasks except Vscan and tasks that you start from Datacap Studio.
 - Each log file contains detailed descriptions of the actions that are run by the task profile.
 - Logging is useful for application troubleshooting.

[Debug and Test](#)

© Copyright IBM Corporation 2017

Figure 7-7. The RRS Log Files

Rulerunner Service (RRS) log files

As Rulerunner runs each action, it writes detailed logging information to a Rulerunner Service (RRS) log file (`task_rrs.log`). Rulerunner also generates an RRS log file whenever you run a task from Datacap Studio.

1. If you want to generate an RRS log file for tasks that you run from the Datacap Web Client or for Datacap Desktop tasks, complete the following steps. Start Datacap Rulerunner Manager.
2. Click the Logging tab.
3. Click the RRS log tab and select the logging options that you want.

In the Datacap Web Client, each task generates its own Rulerunner Service log file. The most recent TravelDocs batches folder contains a log file for each of the task profiles in the Main Job workflow.

Each log file contains detailed descriptions of the actions that are run by the task profile and is useful for application troubleshooting.

Web Client and Datacap Desktop Logging

- Logging of tasks run from either the Datacap Web Client, or Datacap Desktop can be configured
- To log from either or both clients, complete the following steps:
 - Start Datacap Rulerunner Manager.
 - Click the Logging tab.
 - Click the RRS log tab and select the logging options that you want.
- In the web client, each task generates its own Rulerunner Service log file.
- Each log file contains detailed descriptions of the actions run by the task profile.
 - Useful for application troubleshooting.
 - There are good examples in Datacap 9.0 help.

[Debug and Test](#)

© Copyright IBM Corporation 2017

Figure 7-8. Web Client and Datacap Desktop Logging

Enable logging for Datacap Web Client tasks

To enable logging for a web task, you must configure that task in the Datacap Web Client.

1. To enable logging from web client tasks: In the Datacap Web Client, click the Administrator tab.
2. On the Administrator tab, click Workflow.
3. Expand the job that contains the task for which you want to enable logging, and select the task.
4. Click Setup in the Selected task details pane.
5. In the Rulerunner service log field, enter one of the values, as required. Tip: RRS logging is only useful for tasks that run rules. If your web client is not associated with a task profile, an RRS log file is not generated.

In most situations, a setting of 3 provides enough information to help you debug rule-related issues.

Setting Rulerunner Logging by Application/Task

- To configure Rulerunner logging by application and task, string values must be added to the Registry.
- Steps to enable RRS logging by application name:
 - Add a String Value with the Value name `app_filter` to the registry.
 - Then enter the task name in the Value Data field.
 - For example, if you want to enable logging for the APT application tasks, add the `app_filter` value name and enter APT in the Value Data field.
- You can also enable RRS logging by task profile.
- Detailed information is in the Datacap 9.0 online help.

Figure 7-9. Setting Rulerunner Logging by Application/Task

Set Rulerunner logging by application and task

You can set Rulerunner logging by application and by task when you add string values to the registry.

To enable RRS logging by application name, add a String Value with the Value name `app_filter` to the registry. Then, enter the application name in the Value data field. For example, if you want to enable logging for the APT application tasks, add the `app_filter` value name and enter APT in the Value data field.

- To enable RRS logging by task profile, add a String Value with the Value name `tprofile_filter` to the registry. Then, enter the task name in the Value data field. For example, if you want to enable logging for the VScan task, add the `tprofile_filter` value name and enter VScan in the Value data field.
 - For 32-bit (x86) OS, add the value in the following location.
 - `HKEY_LOCAL_MACHINE\SOFTWARE\Datcap\Rulerunner\RRSLog`
 - For 64-bit (x64) OS, add the value in the following location.
 - `HKEY_LOCAL_MACHINE\SOFTWARE\Wow6432Node\Datcap\Rulerunner\RRSLog`

You can set logging for an application and a task profile by adding the app_filter and the tprofile_filter string values to the registry. For example, if you set the app_filter value to APT and the tprofile_filter value to VScan, RRS logging is enabled only when the VScan task profile in the APT application is run.

If the app_filter Value data field is empty, RRS logging is enabled for all applications. If the tprofile_filter Value data field is empty, RRS logging is enabled for all task profiles.

- To confirm that the registry log filter is enabled, look in the Rulerunner.log file or look in the batch. The Rulerunner.log file indicates if application or task filtering is enabled for RRS logging.
- The batch has only the .rrs log files for the filtered task. For example, if you set the tprofile_filter value to VScan, then only the VScan.rrs logs are in the batch.

Note: You must restart the Rulerunner Service when you add a registry key log filter or change the value of a log filter.

Datacap Web Client Tasks and Log Files

- Logging is enabled by default for all tasks except Vscan and tasks that you start from Datacap Studio.
- If you want to generate a task log for Datacap Web Client tasks, you must set the severity level to 1 or higher (on a scale of 0-8).
 - The severity level and the options in the Datacap Web Client Task setup window determine how much information is written to the log file.
- The Task Log file is saved in the batch folder.
 - It provides information that is typically most helpful to IBM support because it documents mostly internal calls.

Figure 7-10. Datacap Web Client Tasks and Log Files

Support Recommendations

- One of the best ways to debug an application is to read the relevant log files after asking yourself “What is failing?”
 - Become familiar with the information that is in the various log files so that when you run into a problem, you’re already familiar with the contents of the files and the kinds of information they contain for non-problematic batches.
 - When a problem occurs, ask things like whether it fails running manually? Is it working outside of Rulerunner server? Are you using a client application (ie. DotScan)? Are there issues with TM server? If so, gather the tmserver log.
- Be aware of what is logged in the Rulerunner log file.
 - Also be aware of the information in the rrs log for ie. a correctly run Page ID step in your workflow.
 - That will be helpful so that you recognize lines of information that might be problem related when Page ID has an error.

Further Recommendations

- Before you run into problems, make sure you know how to configure and use the debugging tools.
- Search in the extensive help which ships with the product.
- A couple of extremely helpful searches related to logs are the following:
 - <http://www.ibm.com/Search/?q=datacap+logging&v=17&en=utf&lang=en&cc=us>
 - http://www-01.ibm.com/support/knowledgecenter/SSZRWV_8.1.0/com.ibm.dc.admin.doc/dclog005.htm
- Notice that there is a reference to the Datacap 8.1 knowledge center.
 - Much of Datacap 8.1 help is relevant to Datacap 9.0.
 - There is much more 8.1 debug information because Datacap 9.0 is so new.

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-12. Further Recommendations

Another Key Way to Debug

- Another excellent way to debug applications is to use the Test Tab in Datacap Studio – covered in last lesson.
- An incomplete list of things you can do using the Test Tab:
 - Set breakpoints.
 - Breakpoints stop a task at a predetermined rule set, rule, or action.
 - You can also configure a breakpoint that stops a task when it starts processing a specific document, page, or field.
 - Single step through your code.
 - Useful to determine whether the functions and actions within a rule are operating as intended.
 - As you step through each line, you can see the actions that are returned as True (check mark) and False (exclamation point).
 - View output that is written to different log files via the Output tab in Datacap Studio.

Figure 7-13. Another Key Way to Debug

Exercise: Application Debugging

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-14. Exercise: Application Debugging

Exercise objectives

- Configure and access logs for debugging

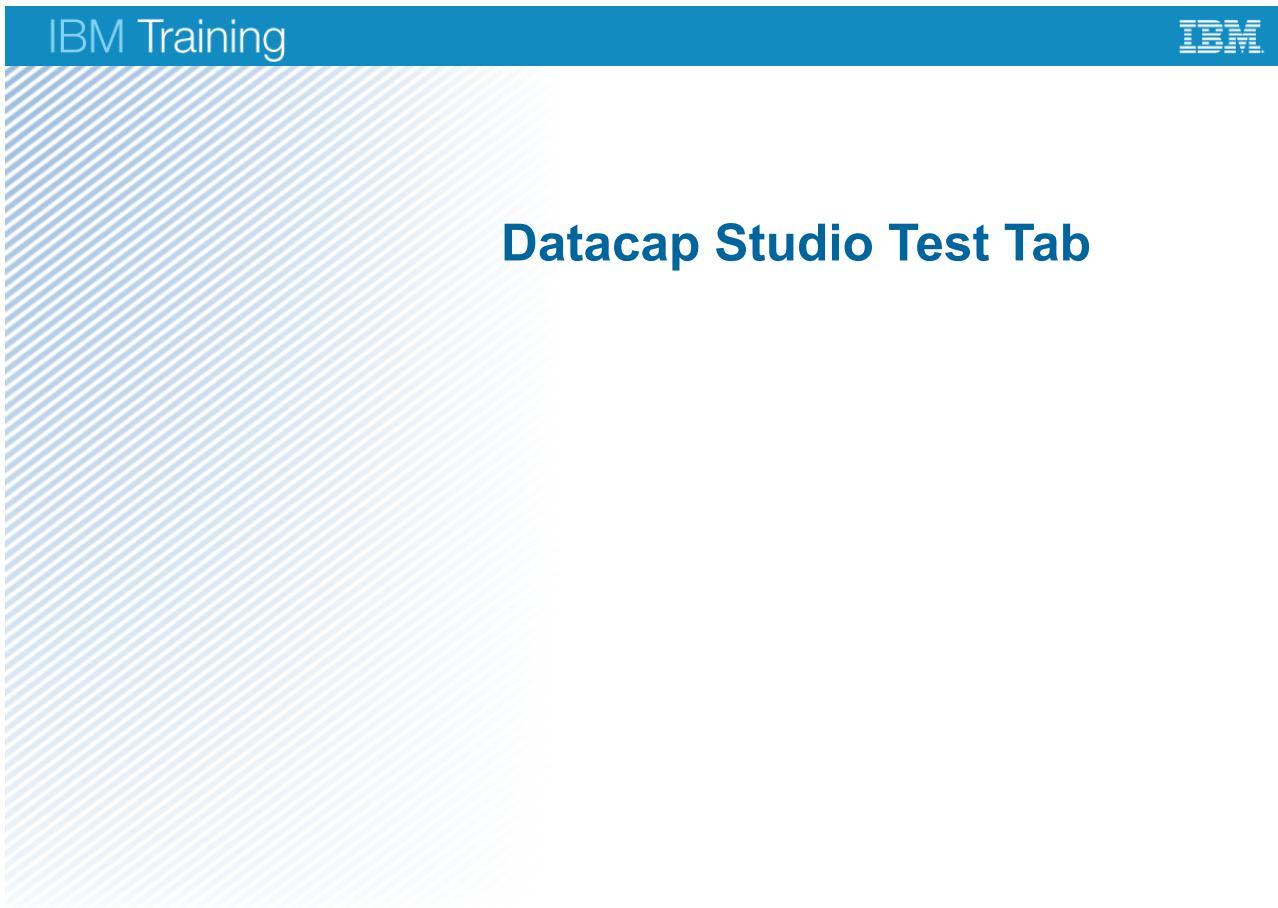


Debug and Test

© Copyright IBM Corporation 2017

Figure 7-15. Exercise objectives

Lesson 7.2. Datacap Studio Test Tab



Debug and Test

© Copyright IBM Corporation 2017

Figure 7-16. Datacap Studio Test Tab



Topics

- Application Debugging
- Datacap Studio Test Tab

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-17. Topics

Why is this lesson important to you?

- As a Datacap business analyst, you build and deploy applications with the Datacap Capture system and communicate solution details to the solution architect, administrator, and users. As part of that, you must be able to debug solutions.
 - To configure and test most Datacap capture applications, it's very helpful, and speeds the development cycle, if the developer makes use of the Datacap Studio Test Tab.
 - The Datacap Studio Test Tab includes debugging features.
 - You can run an application from Datacap Studio to monitor it during execution and determine whether the rules are running as you expect.

Figure 7-18. Why is this lesson important to you?

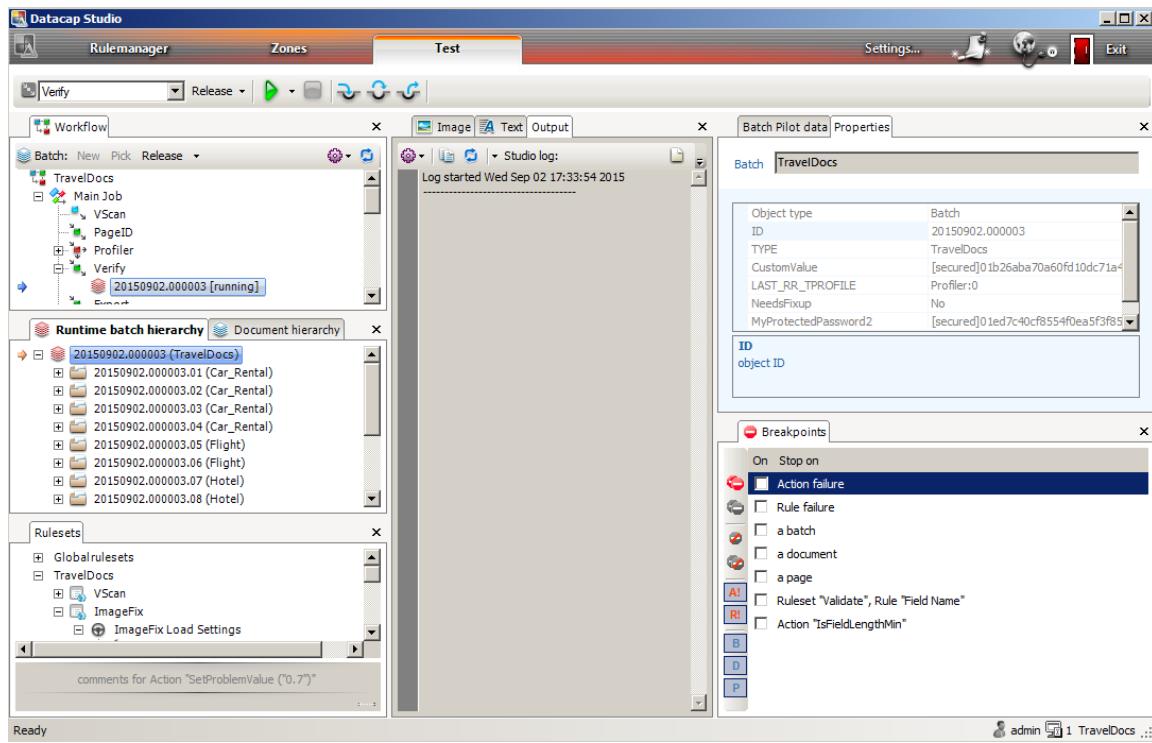
Datacap Studio Test Tab and Debugging

- The Datacap Studio test tab contains 8 panels:
 - Workflow
 - Runtime batch hierarchy
 - Document hierarchy
 - Rulesets
 - Image/Text
 - Batch Data
 - Properties
 - Breakpoints/Runtime state/Call stack

Figure 7-19. Datacap Studio Test Tab and Debugging



Datacap Studio Test Tab



Debug and Test

© Copyright IBM Corporation 2017

Figure 7-20. Datacap Studio Test Tab

The screen capture on the slide illustrates the Test view of Datacap Studio.

- Control Bar
- Workflow pane
- Hierarchy Pane
- Ruleset pane
- Image pane
- BatchPilot data
- Properties
- Breakpoints

Four of the Eight Panels

- Workflow
 - Displays the job types and tasks defined in the Administrator tab.
 - Can run a batch through the workflow.
- Runtime Batch Hierarchy
 - When a batch is running, displays runtime batch hierarchy, including any data values.
 - If you select a page object, the page is displayed in the image panel.
- Document Hierarchy
 - Displays the structure of the documents you are processing.
 - Shows how each element within the structure is processed.
- Rulesets
 - Displays the rules, functions, and actions that make up each ruleset.
 - As you step through the workflow, you can see the current execution point.

Other Four Panels

- Image/Text
 - Displays the selected page in the runtime batch hierarchy.
- Batch Data
 - Displays batch level information for the batch that is running.
- Properties
 - Displays the properties for the selected document hierarchy or ruleset object (read only).
- Breakpoints / Runtime State / Call Stack
 - A breakpoint stops processing at a predetermined ruleset, rule, or action.
 - Or it stops the task when a task starts processing a specific document, page, or field.

Test Tab Debugging Features

- Using Breakpoints
 - A breakpoint stops at a predetermined rule set, rule, or action.
 - It might instead stop a task when it starts processing a specific document, page, or field.
- Single-stepping through your code
 - Single-stepping is useful to determine whether the functions and actions within a rule are operating as intended.
 - As you step through each line, you can see the actions that returned True (check mark) and False (exclamation point).
- Examining log files from the Test tab
 - You can access the Output tab in Datacap Studio to view output written to different log files.

Figure 7-23. Test Tab Debugging Features

Setting and Using Breakpoints

- There are two types of breakpoints, both of which halt execution when Rulerunner encounters the specified element.
 - A breakpoint halts execution when the Rulerunner execution manager encounters the specified element, regardless of context.
 - A full breakpoint halts execution when the Rulerunner execution manager encounters the specified element within the same context.

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-24. Setting and Using Breakpoints

Setting Generic Breakpoints

- Generic Breakpoints: The Datacap Test tab includes two controls that you can select to halt processing when any rule or action fails:

Control	Command	Description
A!	Stop on a failed action	Click this button to add a generic breakpoint that halts processing whenever an action fails
R!	Stop on a failed rule	Click this button to add a generic breakpoint that halts processing whenever a rule fails

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-25. Setting Generic Breakpoints

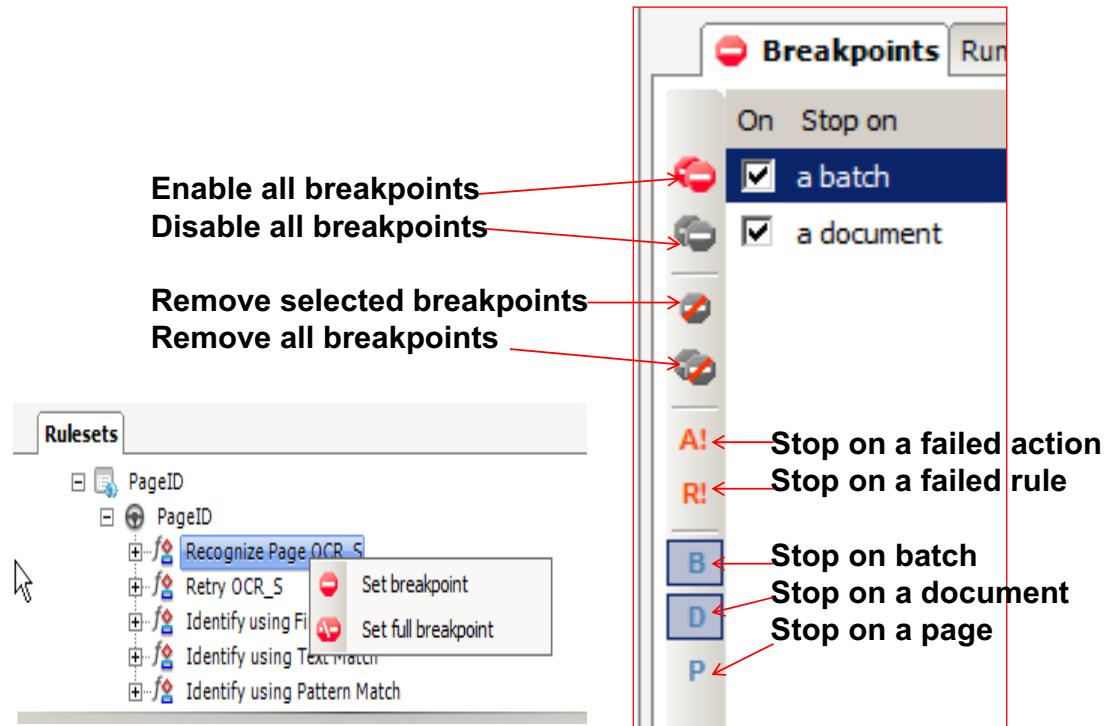
The screen capture on the slide illustrates the generic breakpoint run options:

A! – Stop on a failed action

R! – Stop on a failed rule



Breakpoint Configuration Interfaces



Debug and Test

© Copyright IBM Corporation 2017

Figure 7-26. Breakpoint Configuration Interfaces

The screen capture on the slide illustrates:

- Ruleset pane of the Test tab page
- Breakpoints pane of the Test tab page

Setting Breakpoints

- You can set a breakpoint for a rule set, rule, function, action, document, page, or field.
 - The document, page, or field must exist in the runtime hierarchy before you can set a breakpoint on it.
- To Set Breakpoints:
 - Set a breakpoint on a rule set, rule, function, or action:
 - Go to the rulesets pane on the Datacap Studio Test Tab.
 - Right-click the item and select Set breakpoint or Set full breakpoint.
 - Set a breakpoint on a document, page, or field:
 - Go to the Runtime batch hierarchy pane on the Datacap Studio Test tab.
 - Right-click the item and select Set breakpoint or Set full breakpoint.

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-27. Setting Breakpoints

Disabling and Clearing Breakpoints

- The Breakpoints pane displays all of the defined breakpoints.
- You can enable or disable individual breakpoints by selecting or clearing checkboxes.
 - The buttons on the left of the Breakpoints pane are options to enable, disable, or remove breakpoints.
- The check box to the left of each breakpoint indicates whether the breakpoint is enabled or disabled.
- By default, breakpoints are enabled when you add them.

Figure 7-28. Disabling and Clearing Breakpoints

Single Stepping Through Your Code

- Single stepping is useful to determine whether functions and actions within a rule are operating as intended.
- As you step through each line, you can see the actions that are returned as True (check mark) or False (exclamation point).
 - If an action returns false, you can look in the batch log to see why the action returned false.
- You can also access the Output tab in Datacap Studio to view the output that is written to different log files.

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-29. Single Stepping Through Your Code

How to Step Through Code: Step in

- The Test tab provides UI controls (enabled with tooltips) for stepping through code.
- Step in – steps into the next line of code.
 - If the next line calls a rule or function, Step in opens the rule or function and halts inside it.
 - If the next line is an action, Step in opens the action; you must click it again to close the action.

Figure 7-30. How to Step Through Code: Step in

How to Step Through Code: Step Over or Out

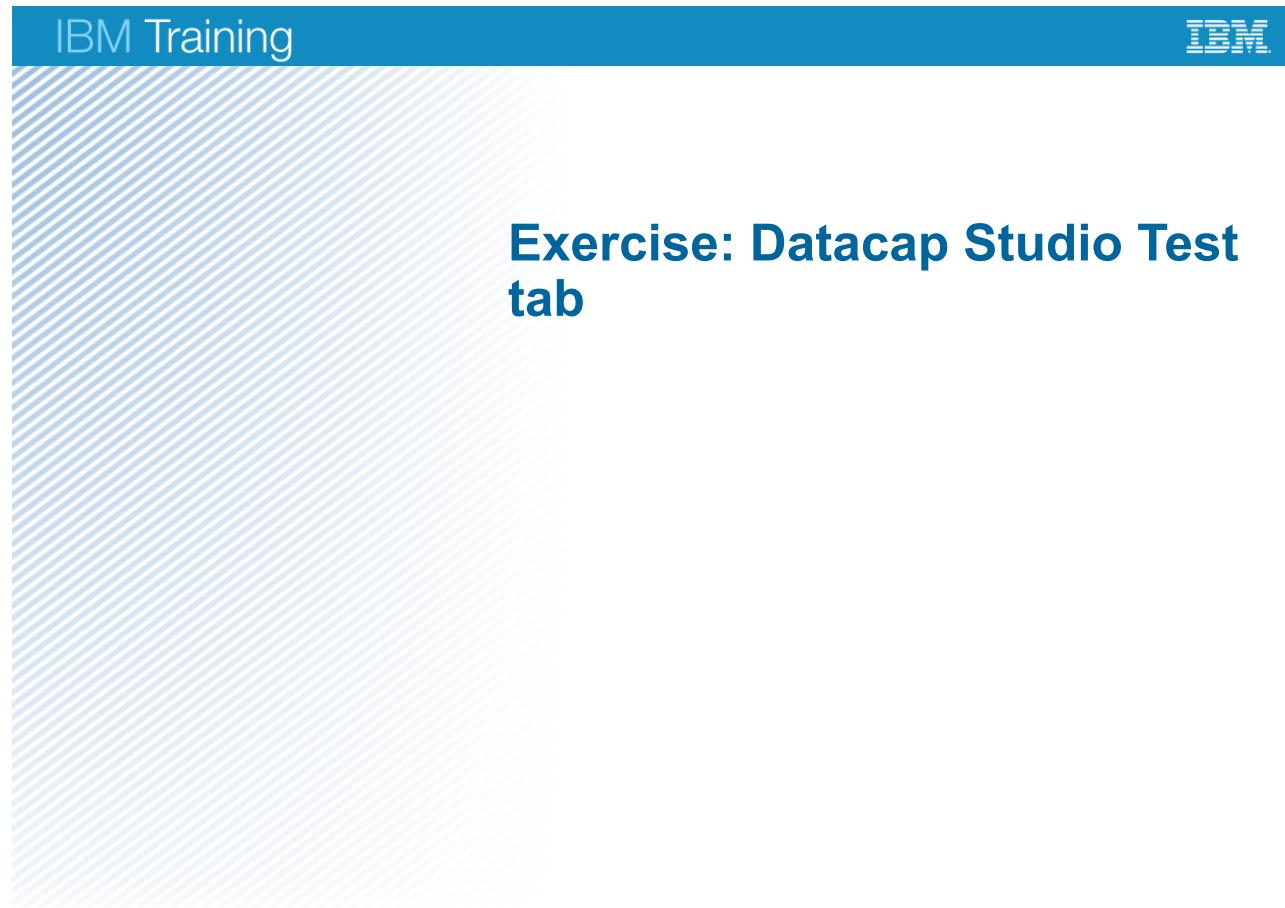
- Step/Step over – starts the next line of code and any lower-level functions and actions, and then stops:
▪ If the next line is an action, Step over works like Step in and opens the action.
- Step out – steps through the next line of code:
▪ If the next line is a rule or function, Step out works like Step over and starts any lower-level functions and actions.
▪ If the next line is an action, Step out starts and closes the action.

Figure 7-31. How to Step Through Code: Step Over or Out

Examining Log Files from the Test Tab

- You can access the Output tab in Datacap Studio to view the information that is written to different log files.
- 1. Click the Output tab in the center pane of the Test tab.
 2. If Batch Log is not already selected, click the down-arrow and select it from the list of available logs.
- The output pane refreshes automatically when:
 - You stop at a breakpoint
 - You single-step through a line of code
 - When the current task profile completes
- You need to scroll to the bottom each time to see the latest messages.

Figure 7-32. Examining Log Files from the Test Tab



Debug and Test

© Copyright IBM Corporation 2017

Figure 7-33. Exercise: Datacap Studio Test tab

Exercise objectives

- Stepping a batch through a PageID task profile



Debug and Test

© Copyright IBM Corporation 2017

Figure 7-34. Exercise objectives

Unit summary

- After completing this unit, you should be able to:
- Configure and access logs for debugging
- Use the Datacap Studio Test tab to step through a batch for testing

Debug and Test

© Copyright IBM Corporation 2017

Figure 7-35. Unit summary



IBM Training



© Copyright International Business Machines Corporation 2017, 2020.