Course Guide

IBM Datacap 9.0.1: Introduction

Course code F251 ERC 1.0

IBM Training

October 2016 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.

Contents

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:

Daeja™ DB2® FileNet®
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: Introduction

Duration: 1 day

Overview

This course is an entry level class that is a prerequisite for all other IBM Datacap 9.0.1 courses.

You work with a fully functioning IBM Datacap system to practice the skills.

This course introduces the Datacap system components, processes, and architecture. You learn how to use Datacap clients to process document batches.

Audience

- Any one who needs to know the Datacap architecture and Datacap clients.
- Administrators who are responsible for configuring and administering Datacap system
- Application Builders who implement data capture solutions using the IBM Datacap software suite.

Agenda

The following agenda lists the lessons in this course and the approximate times that are planned for these events.

Day 1

(00:15) Welcome

(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)

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

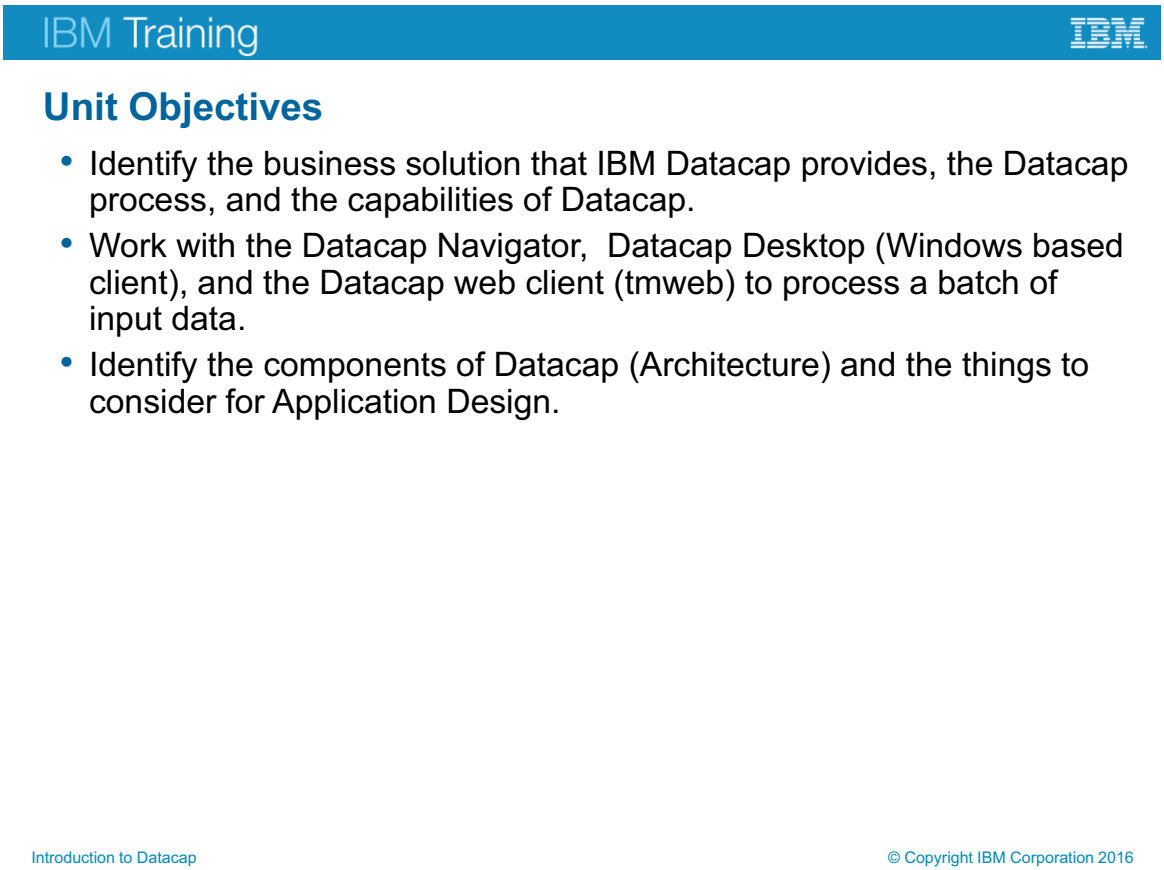


Figure 1-1. Unit Objectives

Lesson 1.1. Datacap overview

Figure 1-2. Datacap overview

Figure 1-3. Lessons

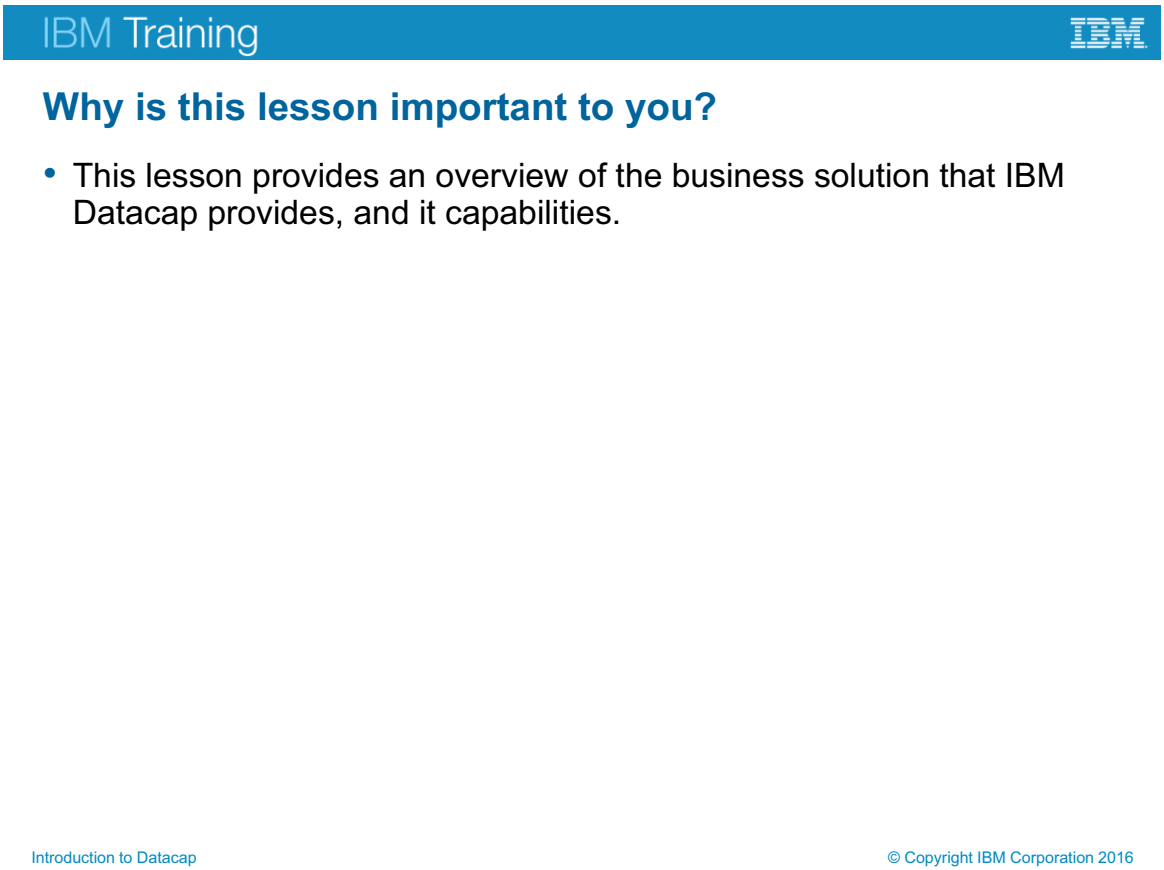


Figure 1-4. Why is this lesson important to you?

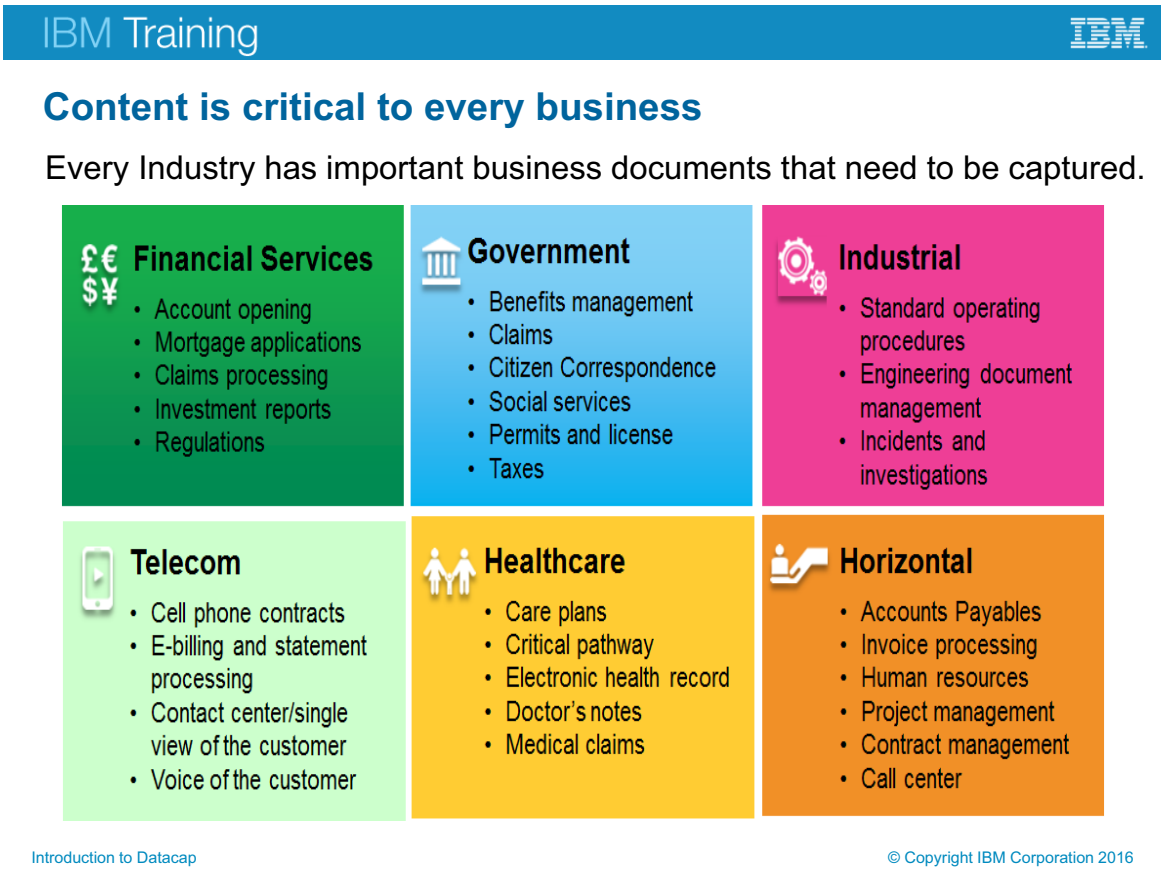


Figure 1-5. Content is critical to every business

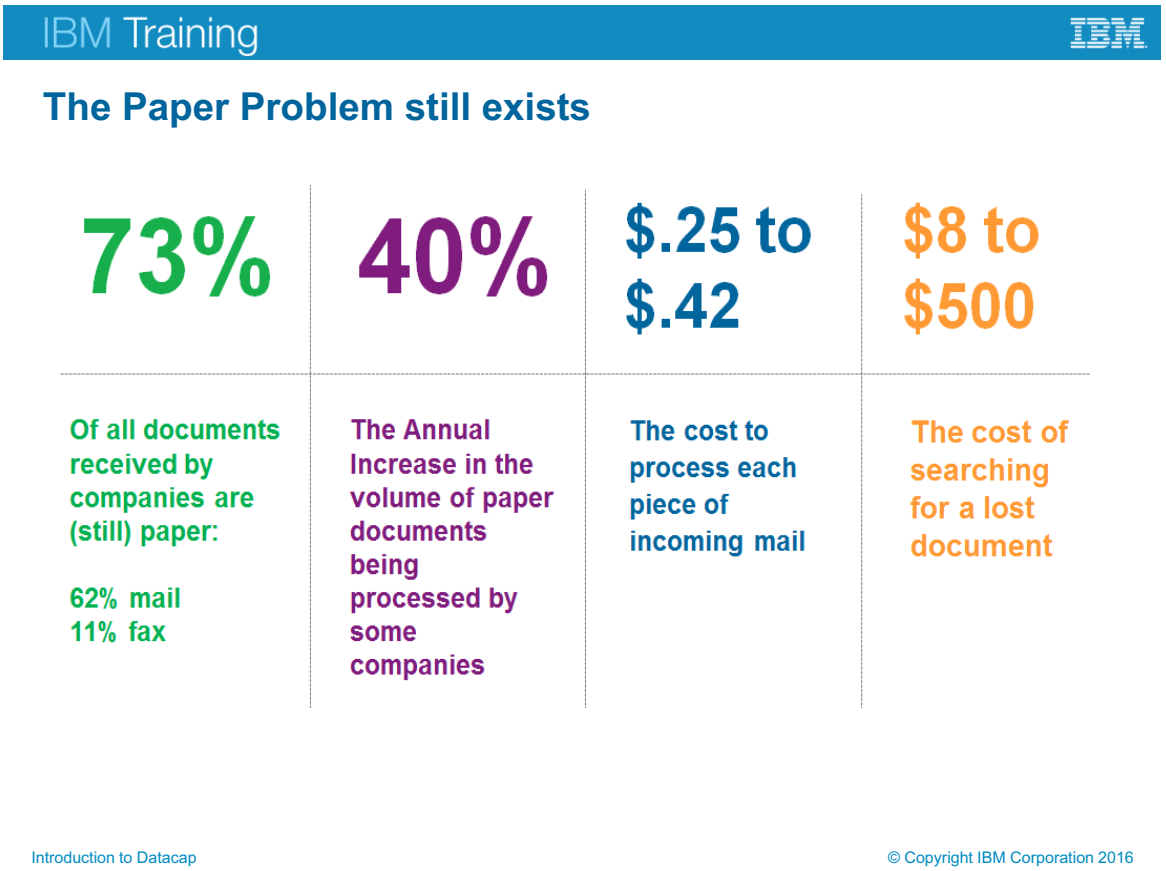


Figure 1-6. The Paper Problem still exists

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.

Figure 1-7. Cost of shipping paper adds up

IBM Training

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

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%

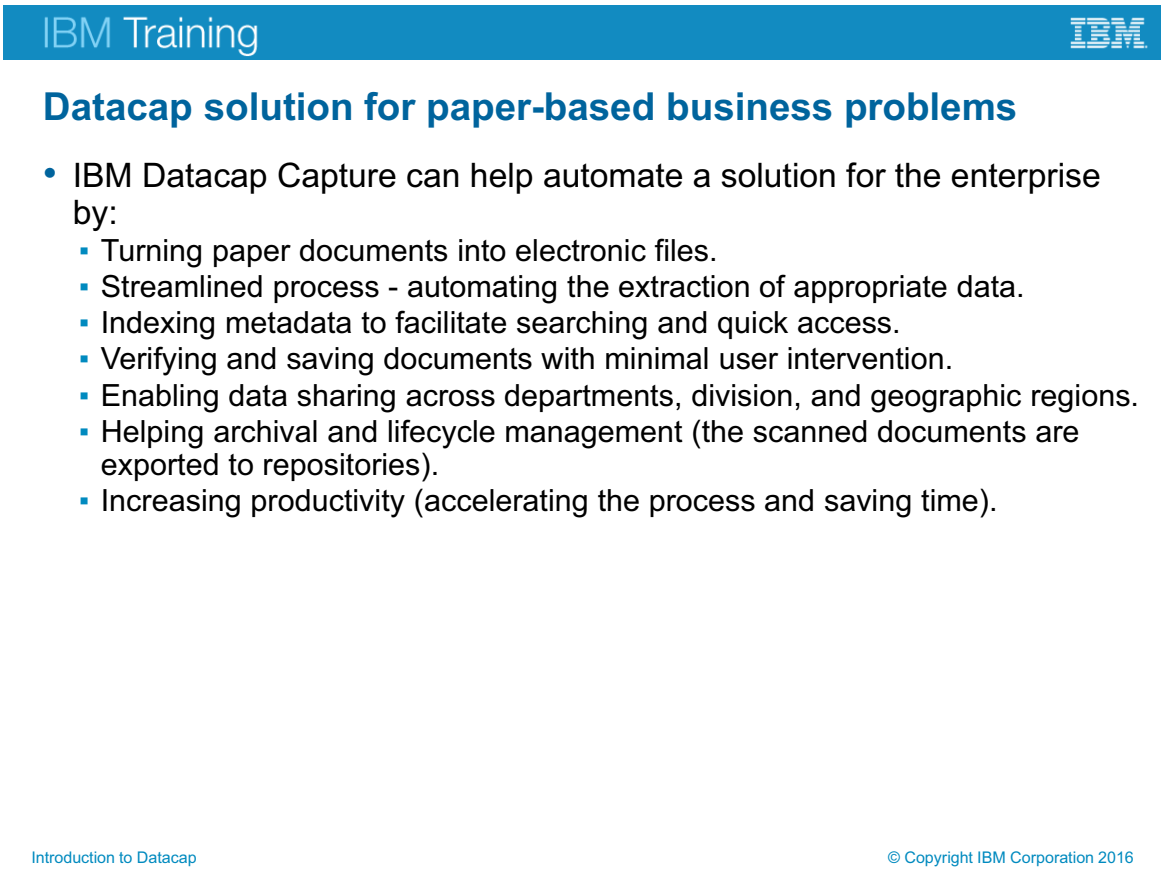


Figure 1-10. Datacap solution for paper-based business problems

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.

IBM Training

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?

IBM Training

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 :	48998813
Incident Number :	CL-4328941
Incident Date:	03/03/11
Incident Time:	14:15
Vehicle License :	2296RG
Vehicle Colour :	Black
Vehicle Manufacturer :	Toa
Vehicle Model :	Quaser
Year of Manufacture :	2009
Chasis 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? ☐ Yes ☒ 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: Chris Jones

Dated: MMDDYY 0 5 0 3 1 1

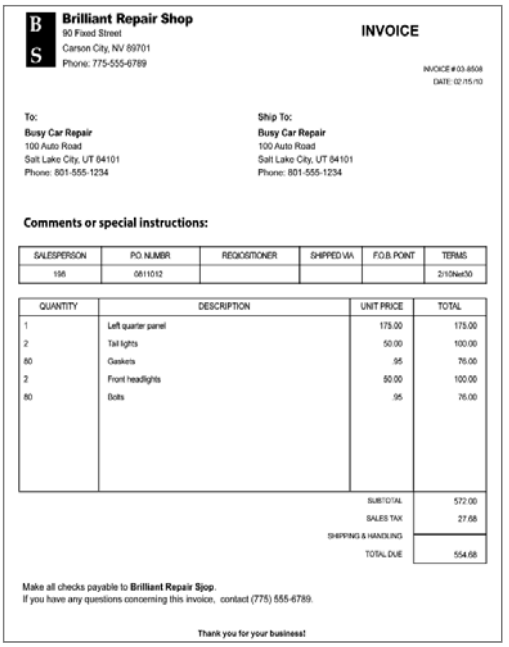
Introduction to Datacap

Figure 1-12. Structured documents

IBM Training

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



Brilliant Repair Shop
90 Fixed Street
Carson City, NV 89701
Phone: 775-555-6789

INVOICE

INVOICE # 03.8508
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	P.O. NUMBER	REQUISITIONER	SHIPPED VIA	F.O.B. POINT	TERMS
196	0811012				2/10Net30

QUANTITY	DESCRIPTION	UNIT PRICE	TOTAL
1	Left quarter panel	175.00	175.00
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	554.68

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!

Figure 1-13. Unstructured documents

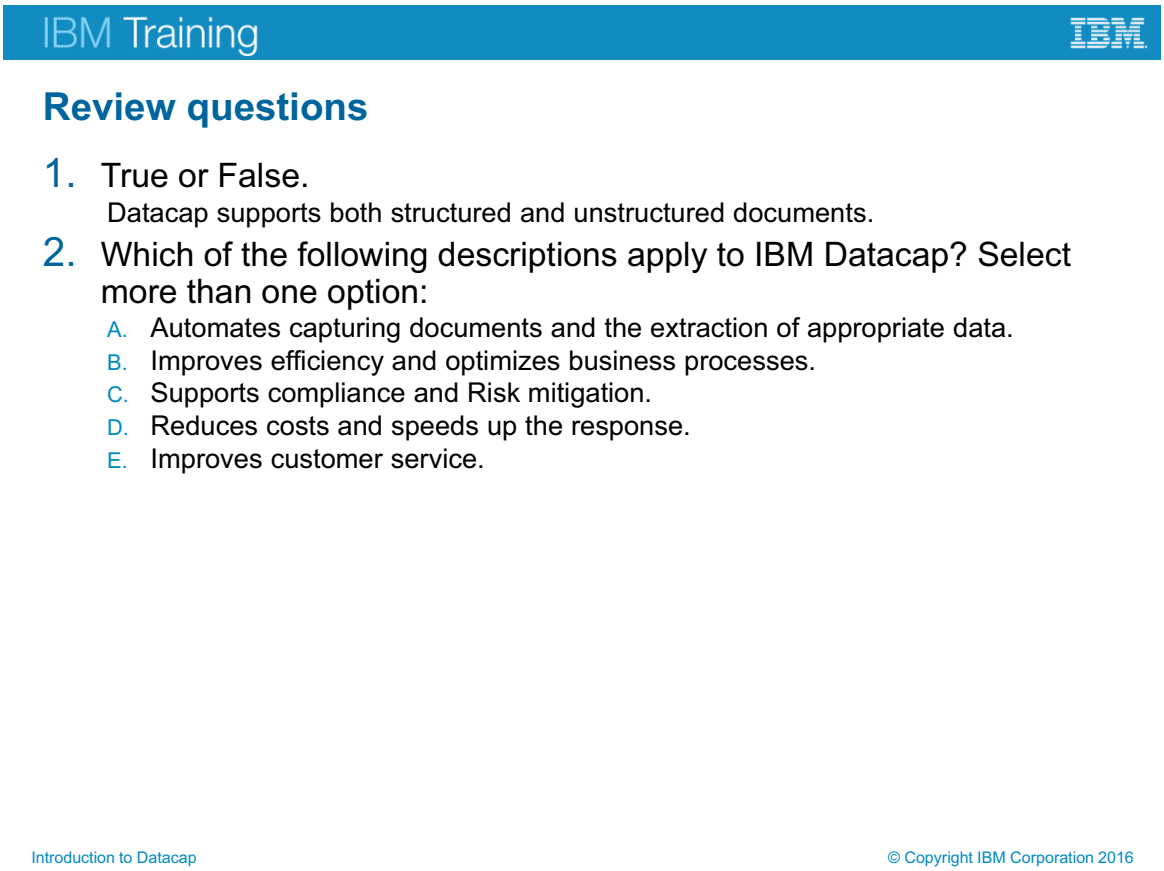


Figure 1-14. Review questions

IBM Training

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

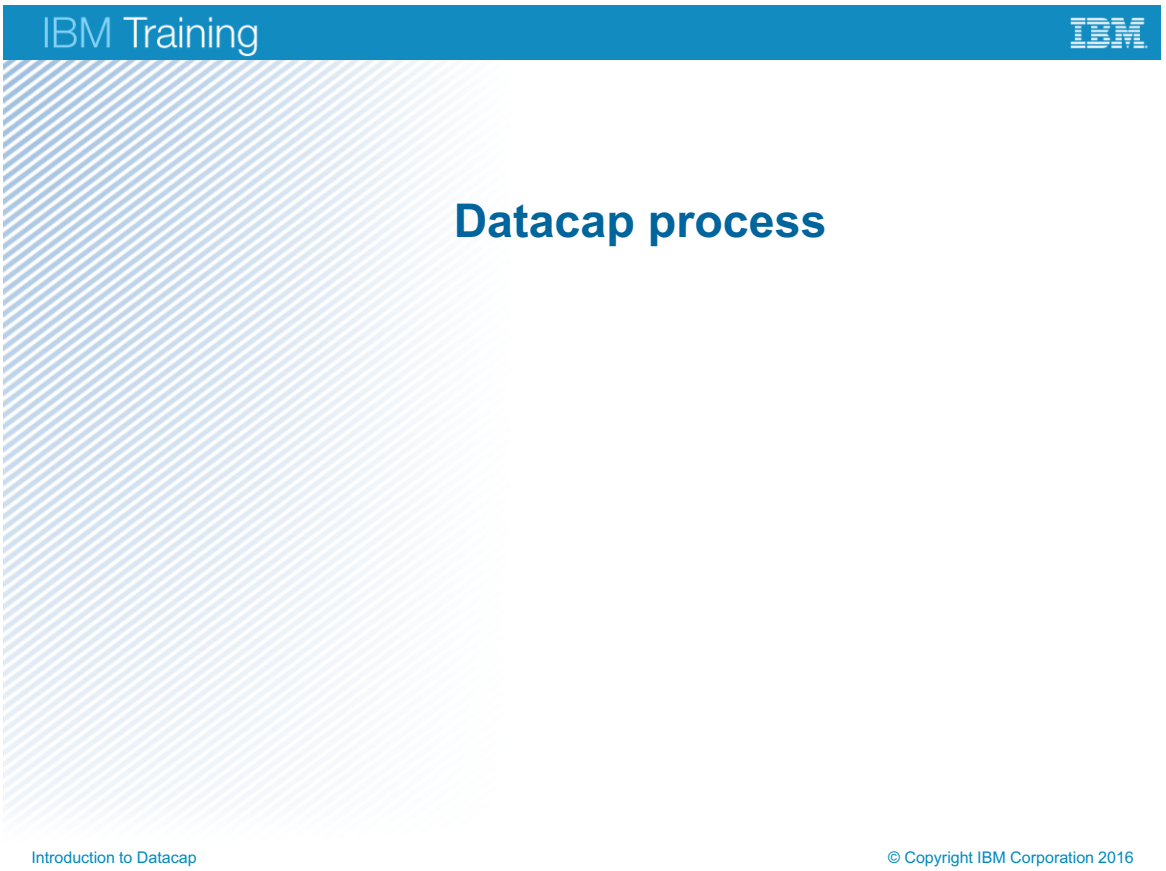


Figure 1-16. Datacap process

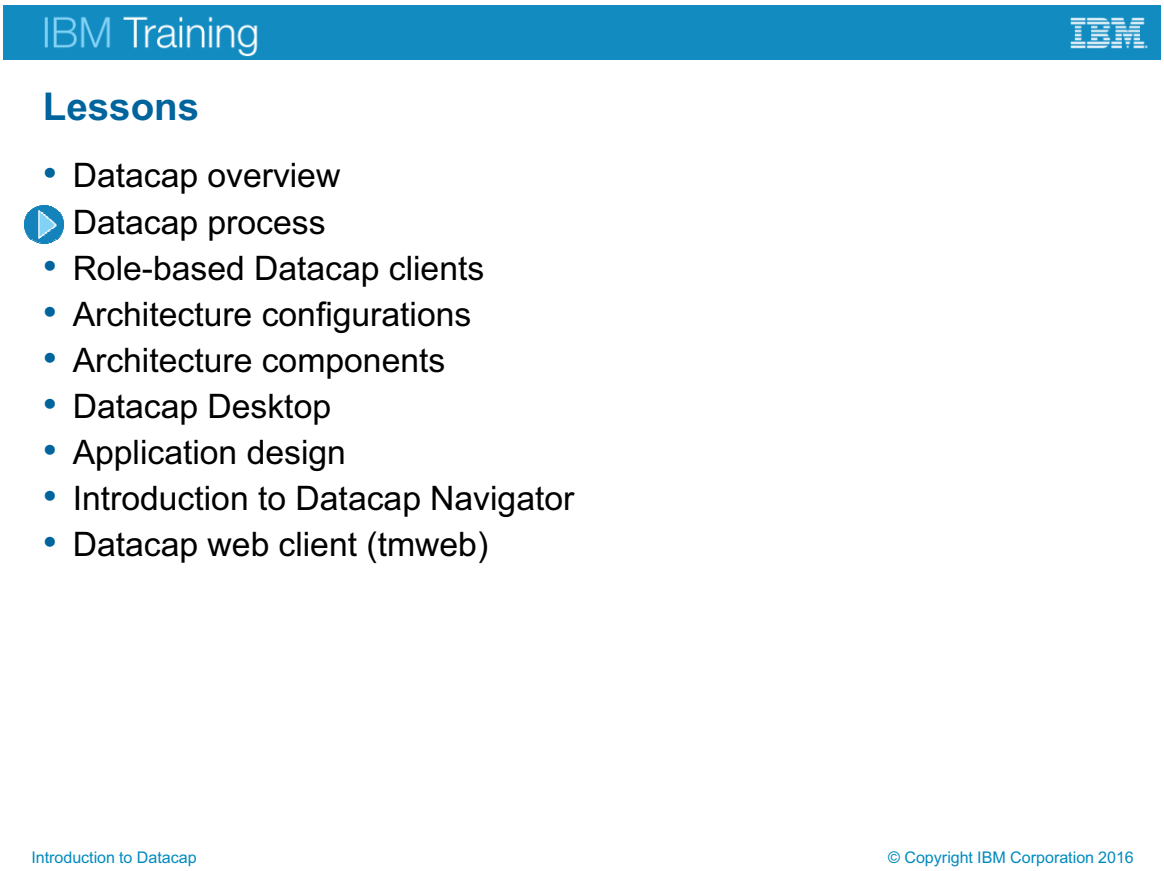


Figure 1-17. Lessons

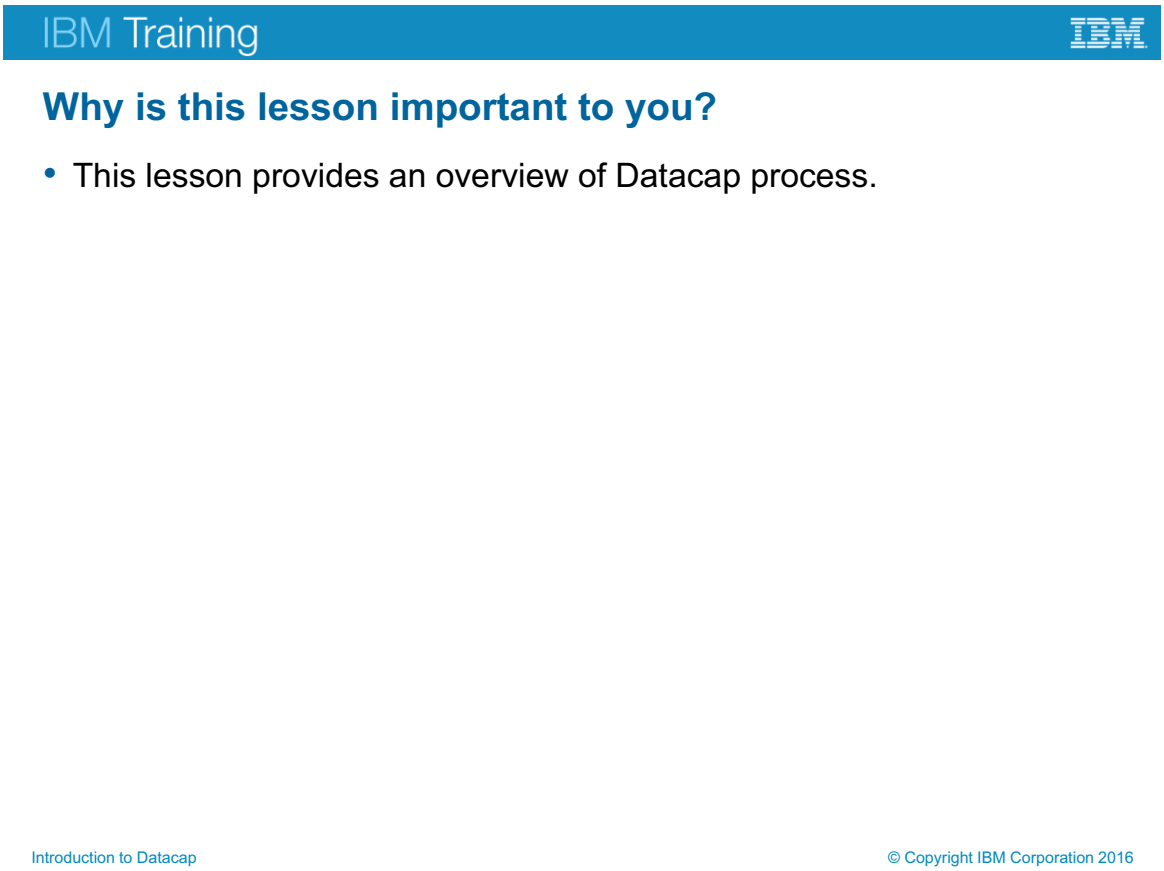


Figure 1-18. Why is this lesson important to you?

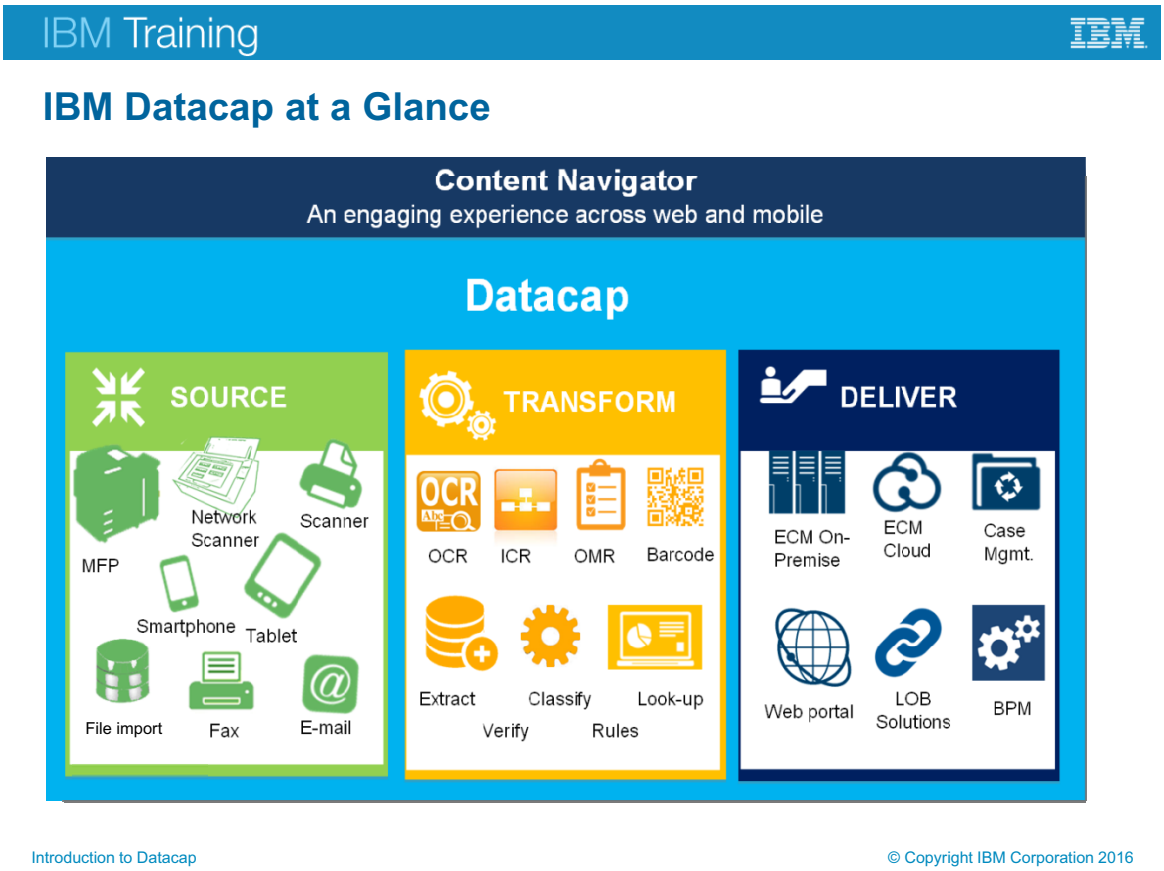


Figure 1-19. IBM Datacap at a Glance

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.

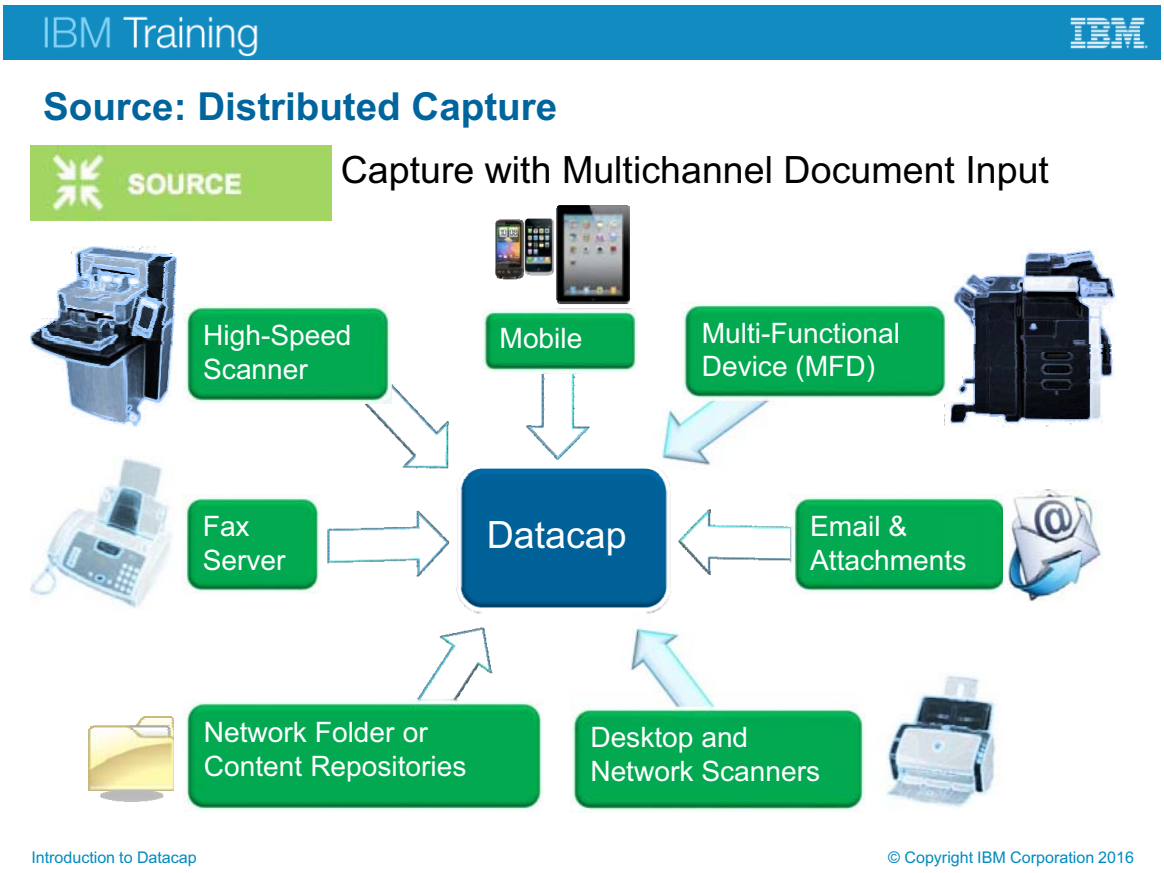


Figure 1-20. Source: Distributed Capture

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

IBM Training

Datacap supported file formats for the source

Input Channels	File Format	Notes
• Scanners • Multi-Functional Devices	• TIFF, JPEG, PDF	Color and grayscale are typically converted to single page bitonal TIFF.
• Mobile devices • Email attachments • Windows file system	• TIFF, JPEG, PDF, PNG • HTML, RTF, TXT, DOC, DOCX, XLS, XLSX, ZIP, EML	• JPEG2000 compression is not supported. • Password-protected PDF files and fillable PDF forms are not supported.
• FAX	• G3, or G4 TIFF • Single-page or multi-page bitonal	• For recognition and all tasks. • Fax documents are black and white.

Introduction to Datacap

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

IBM Training

Transform: Recognition, Classification, and Validation

TRANSFORM

- 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

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.

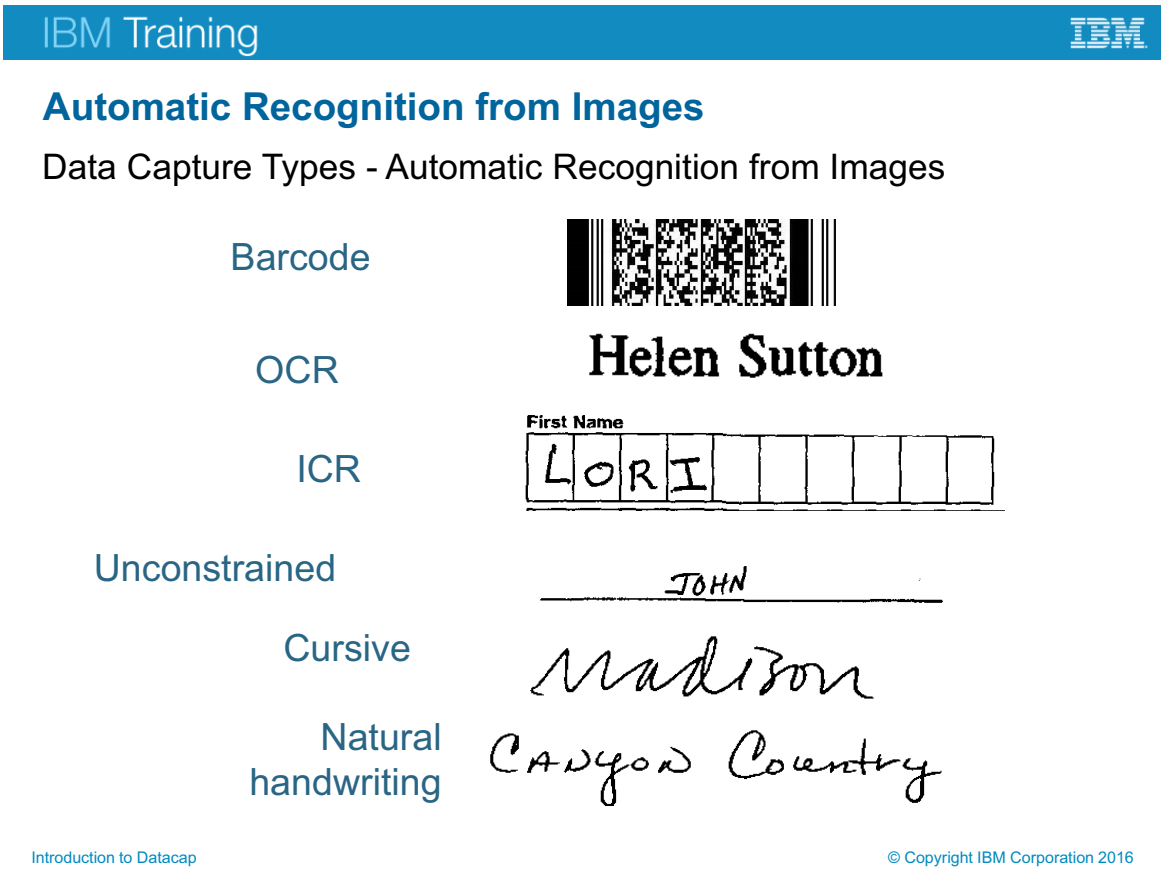


Figure 1-23. Automatic Recognition from Images

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

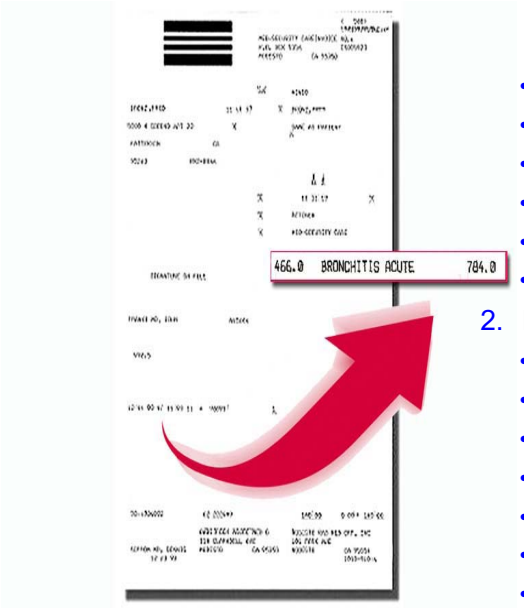
- Recognize check boxes, radio buttons, and so on.

Barcode recognition

- The ability to recognize and interpret barcodes.

IBM Training

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

Figure 1-24. Automatic Data Extraction

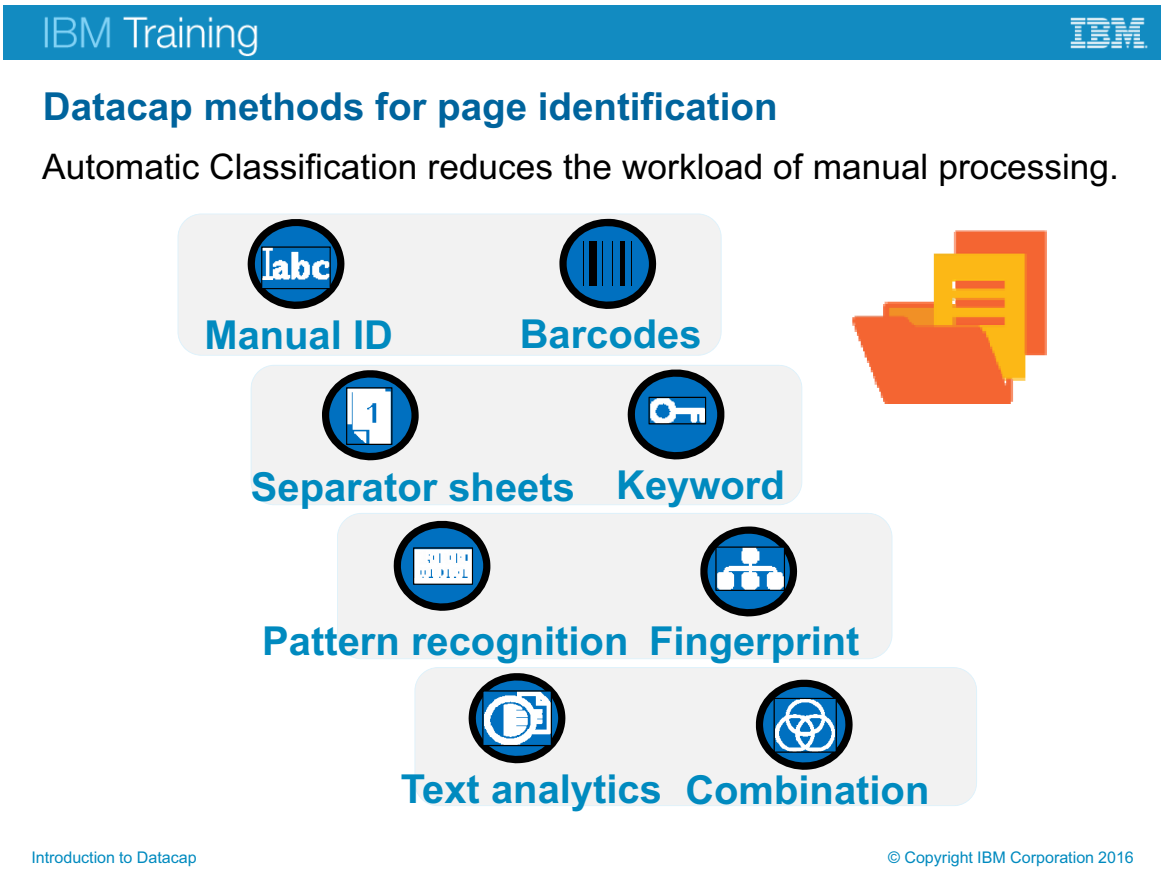


Figure 1-25. Datacap methods for page identification

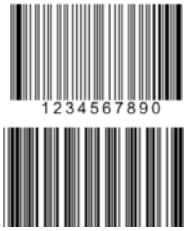
The diagram shows various page identification methods.

Document classification is based on the page identification.

IBM Training

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.

Introduction to Datacap

Figure 1-26. Page identification – Barcodes

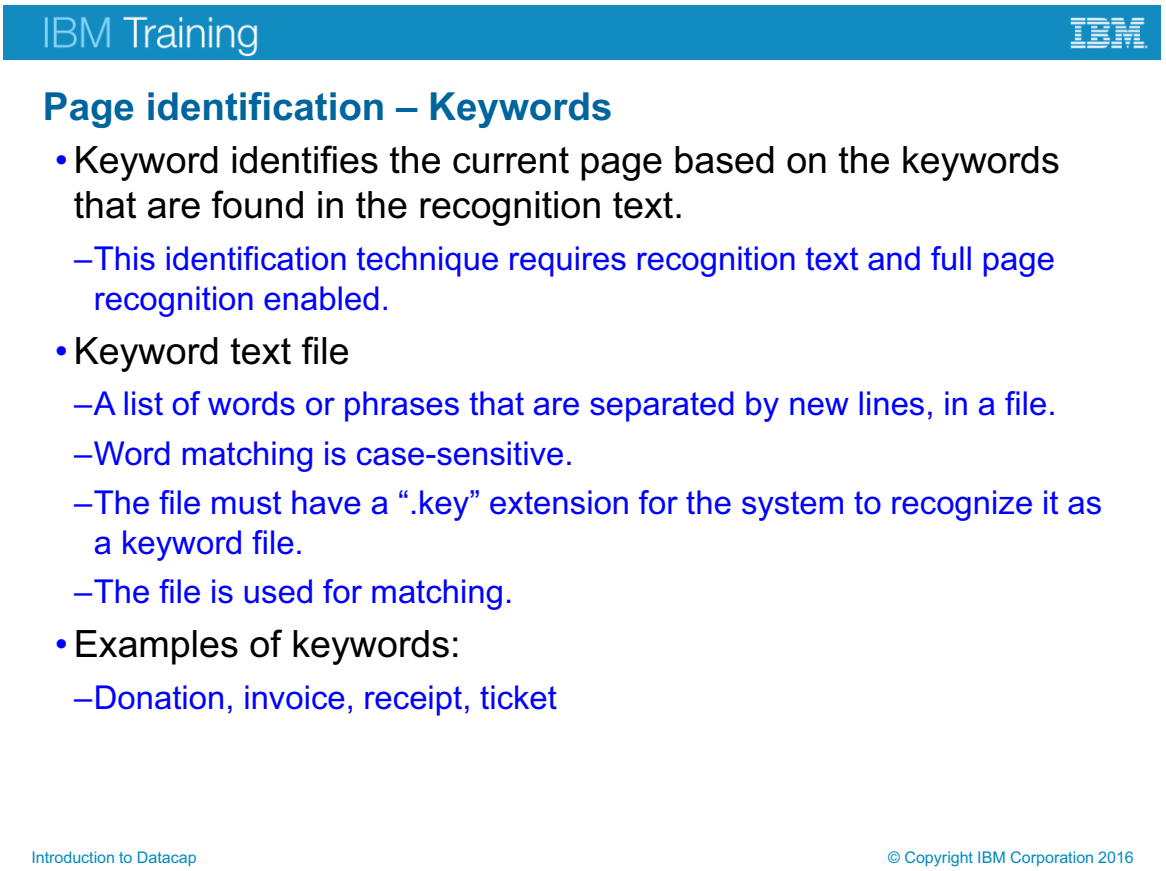


Figure 1-27. Page identification – Keywords

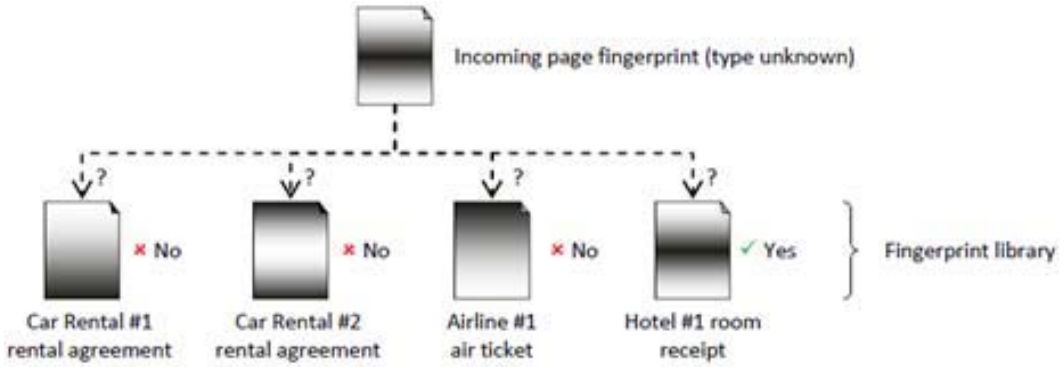
How do keywords work in page identification?

- 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.

IBM Training

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

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

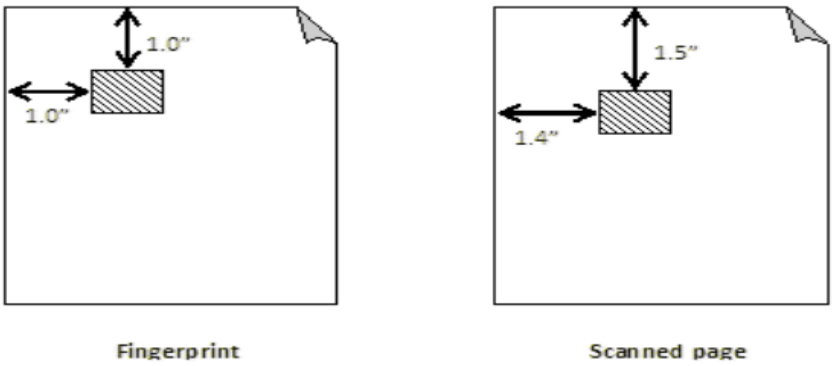
In the example that is shown on the diagram, the incoming page matches the Hotel #1 room receipt.

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

IBM Training

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.



Introduction to Datacap

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.

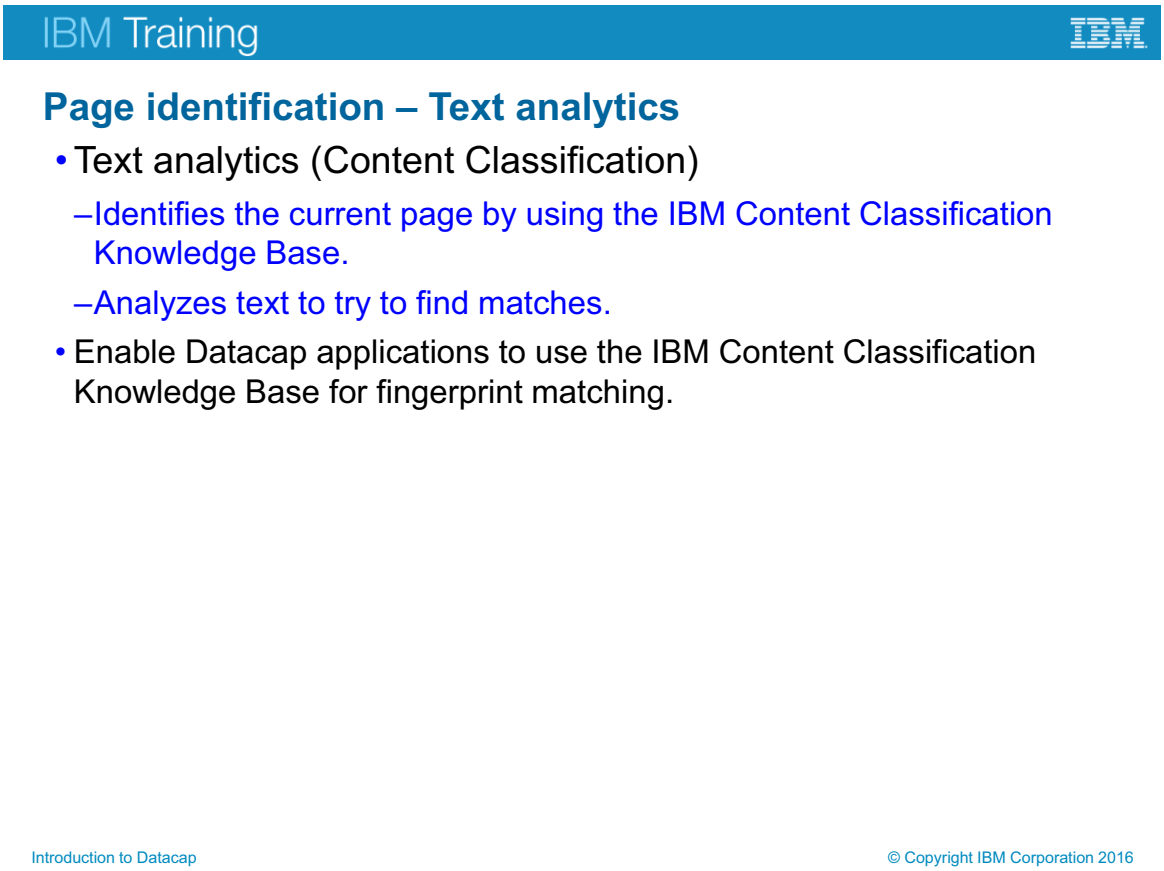


Figure 1-30. Page identification – Text analytics

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.

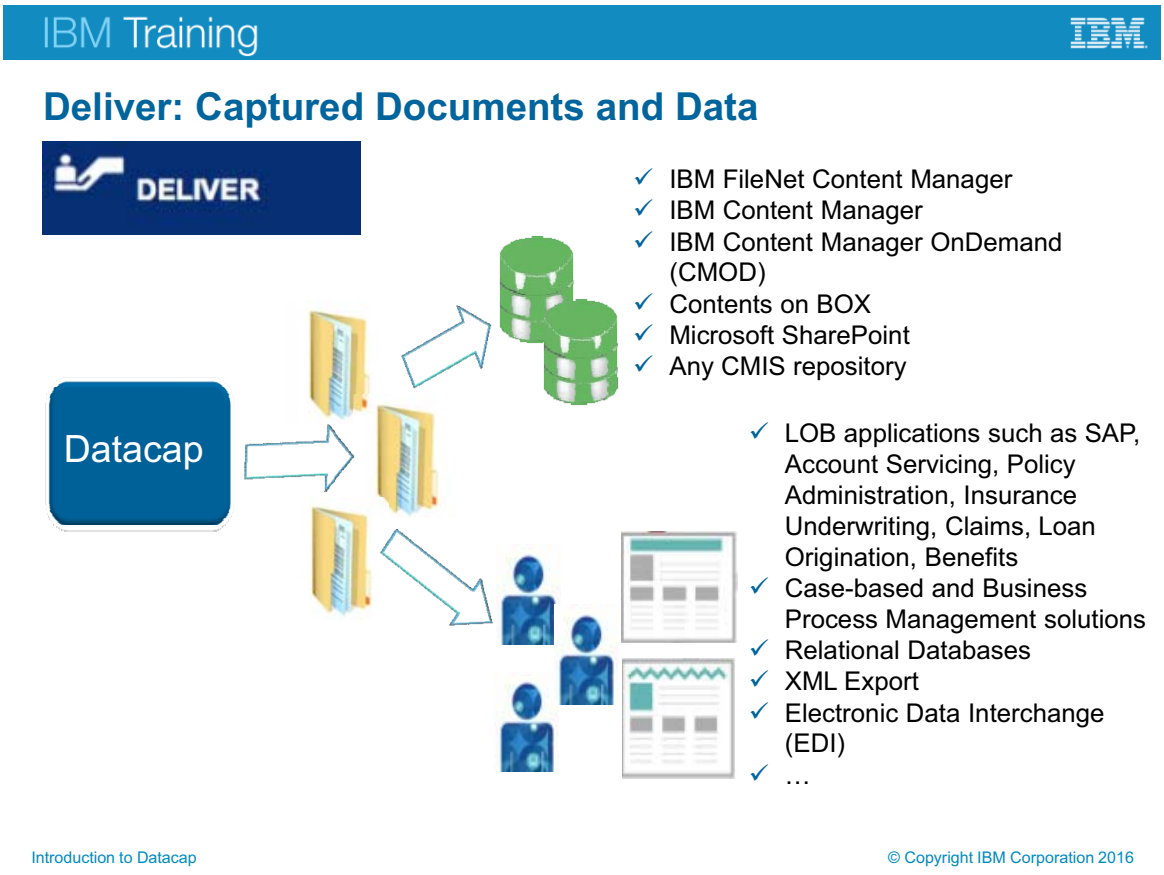


Figure 1-31. Deliver: Captured Documents and Data

Medical Claims Datacap application uses Electronic Data Interchange (EDI)

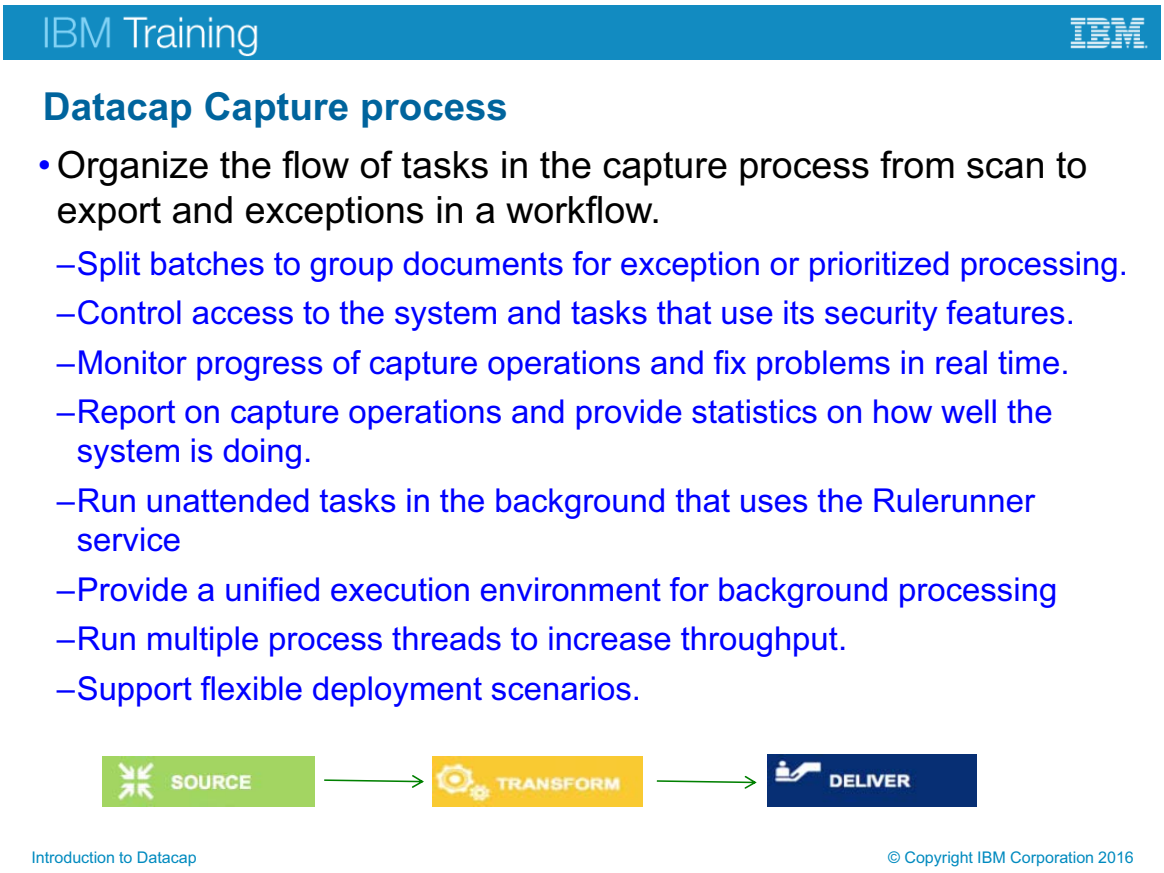


Figure 1-32. Datacap Capture process

Support flexible deployment scenarios

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

IBM Training

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

IBM Training

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

IBM Training

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)

IBM Training

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

Introduction to Datacap

Figure 1-36. Review answers (1)

IBM Training

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)

IBM Training

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

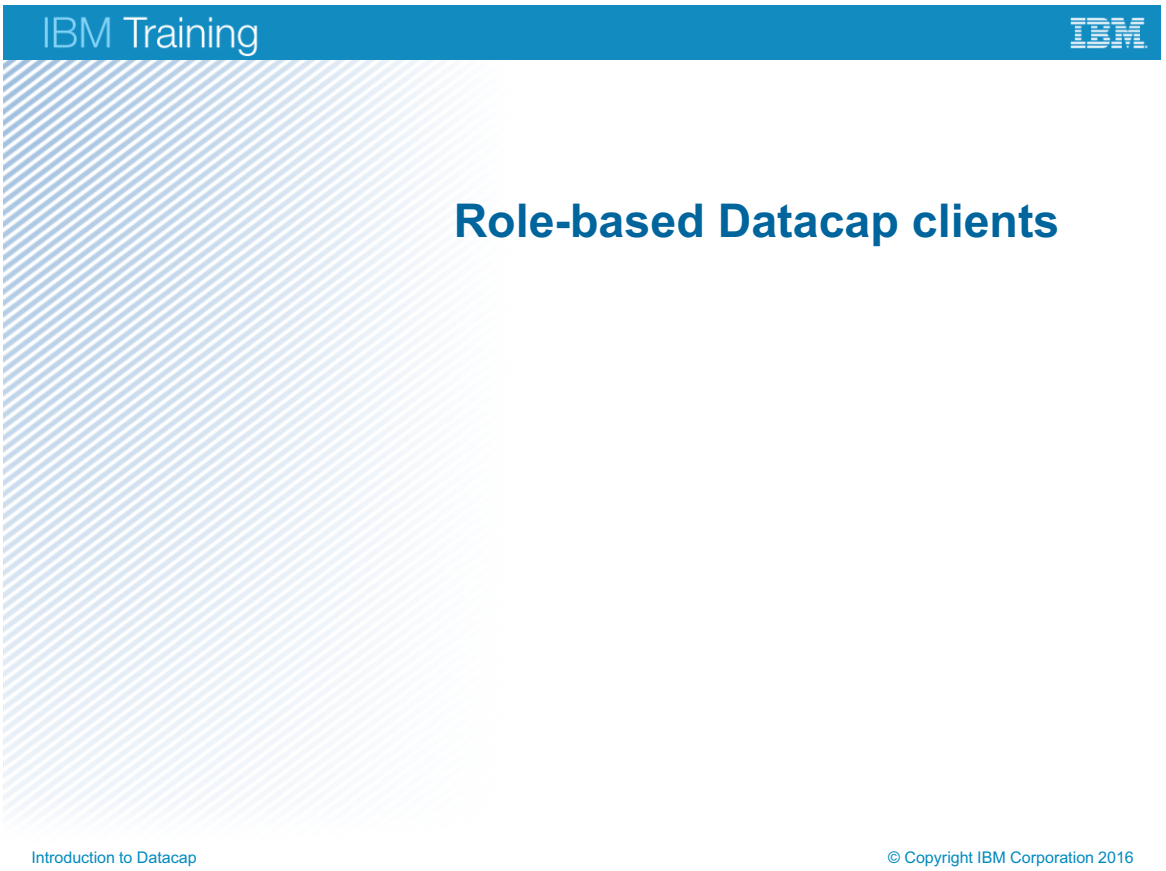


Figure 1-39. Role-based Datacap clients

Figure 1-40. Lessons

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

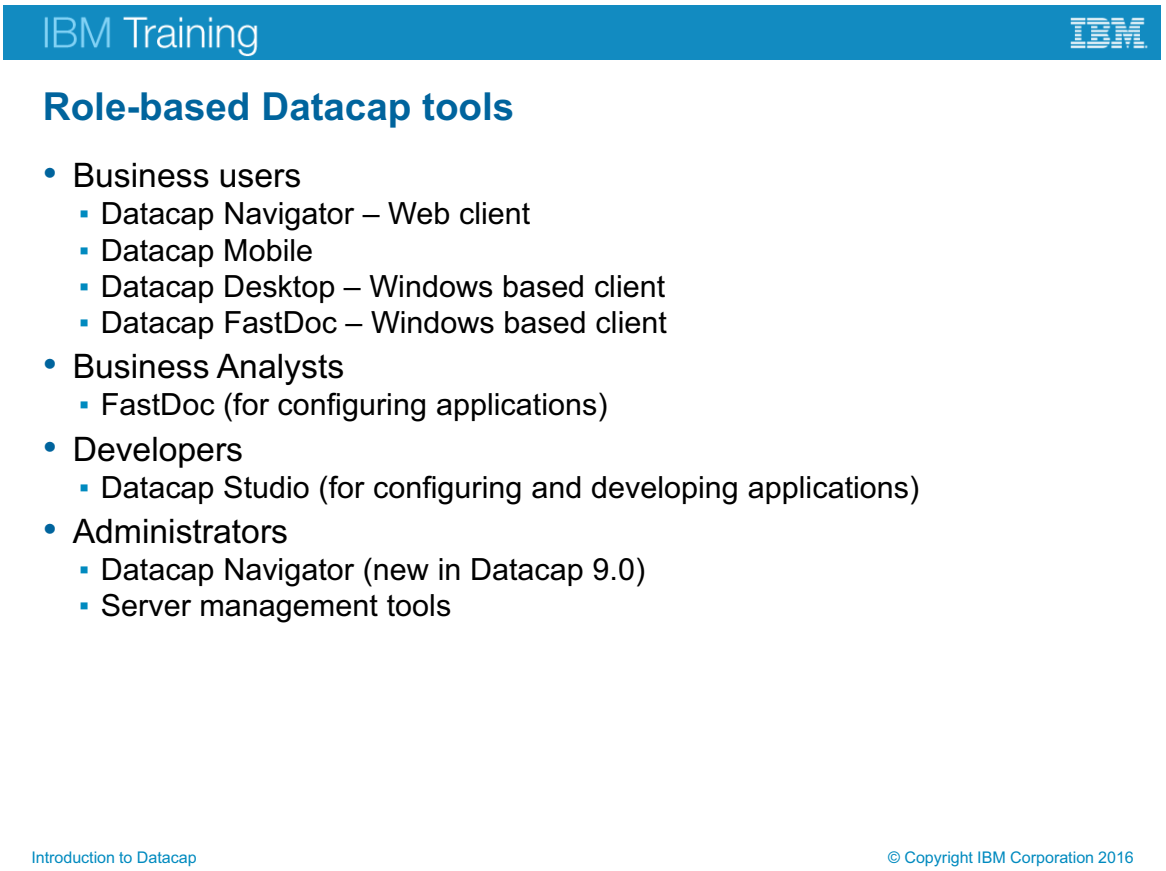


Figure 1-42. Role-based Datacap tools

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.

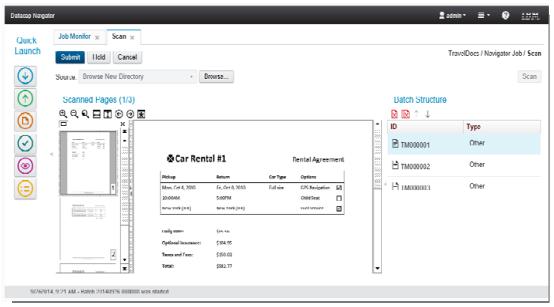
Datacap Studio

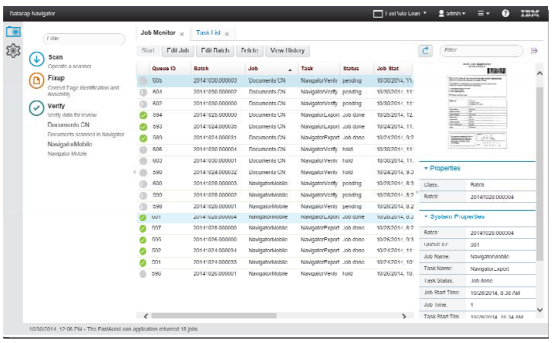
- A more robust development environment to customize FastDoc built applications or build custom applications from scratch.

IBM Training

Datacap Navigator – Web client

Distributed Web Capture with IBM Content Navigator
Scan, Monitor, Verify





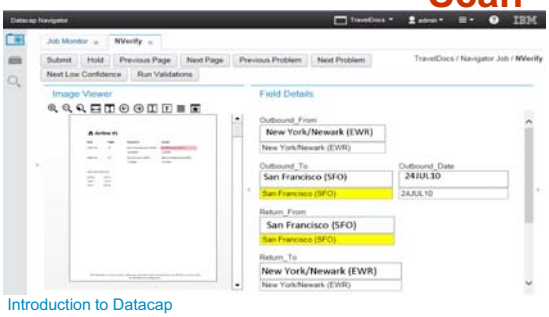


Figure 1-43. Datacap Navigator – Web client

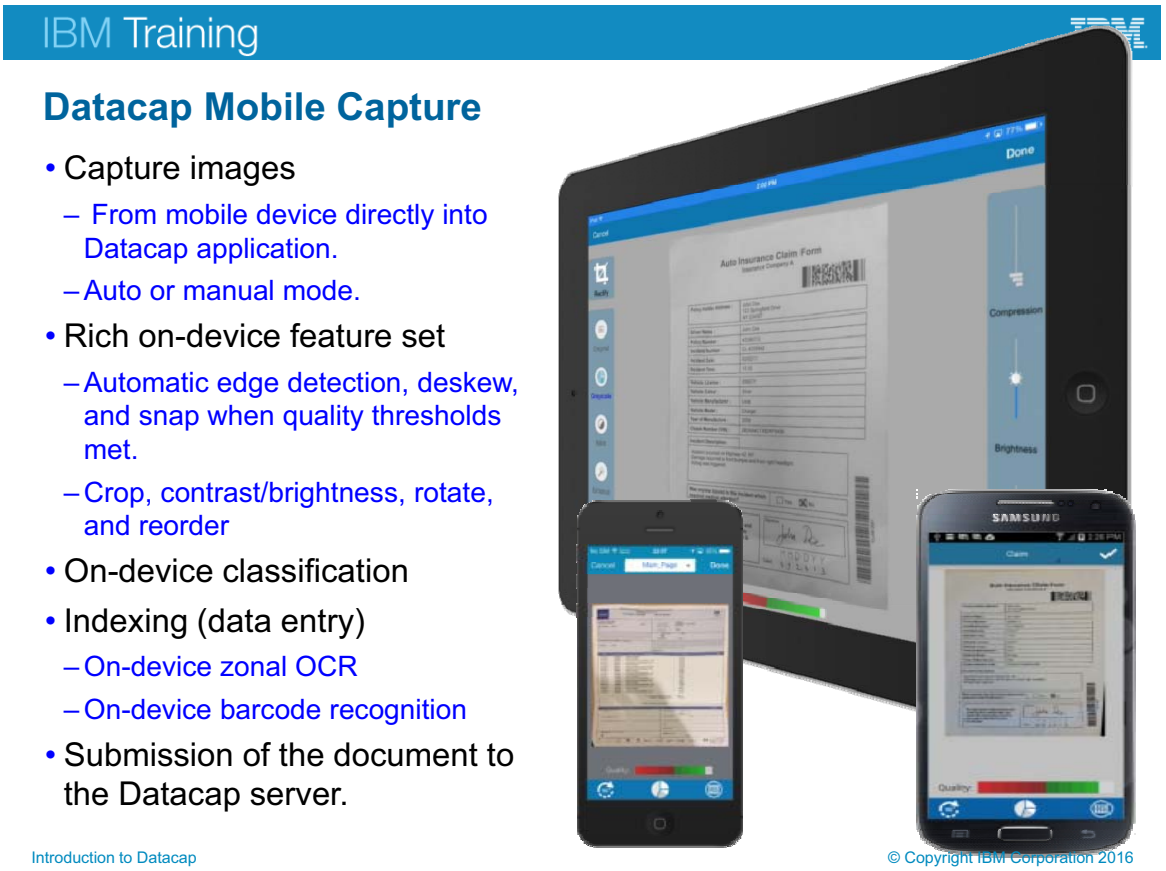


Figure 1-44. Datacap Mobile Capture

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 Training

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

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

IBM Training

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

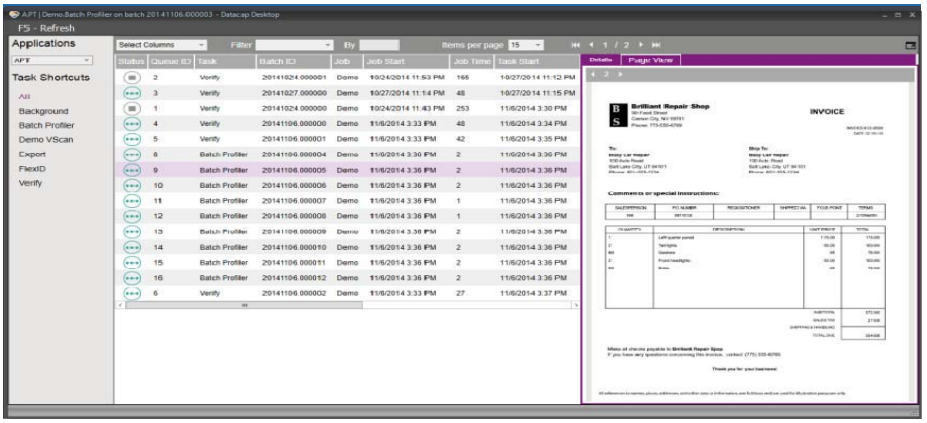


Figure 1-46. Datacap Desktop – Windows based client

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.

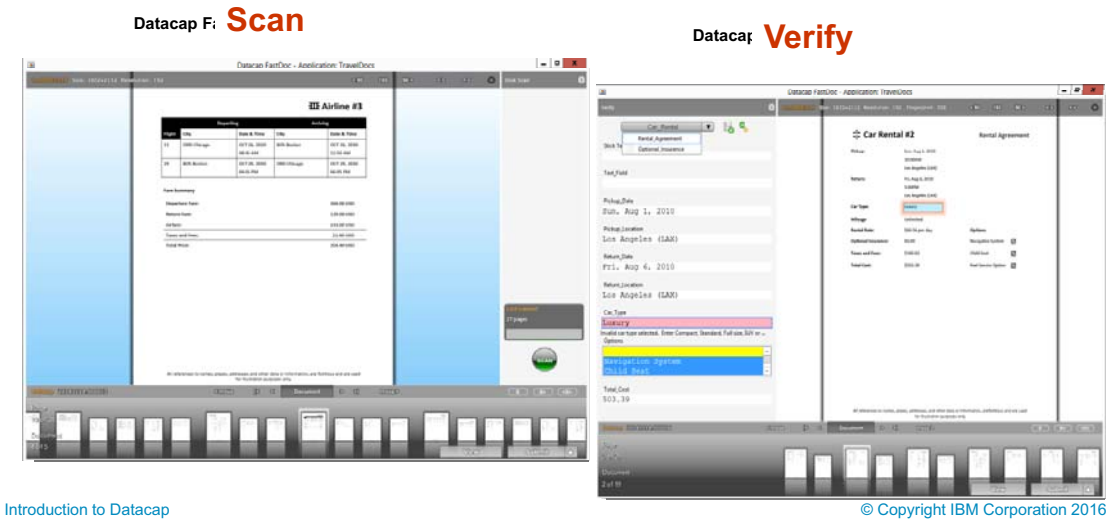


Figure 1-47. Datacap FastDoc – Windows based client

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.

IBM Training

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

IBM Training

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

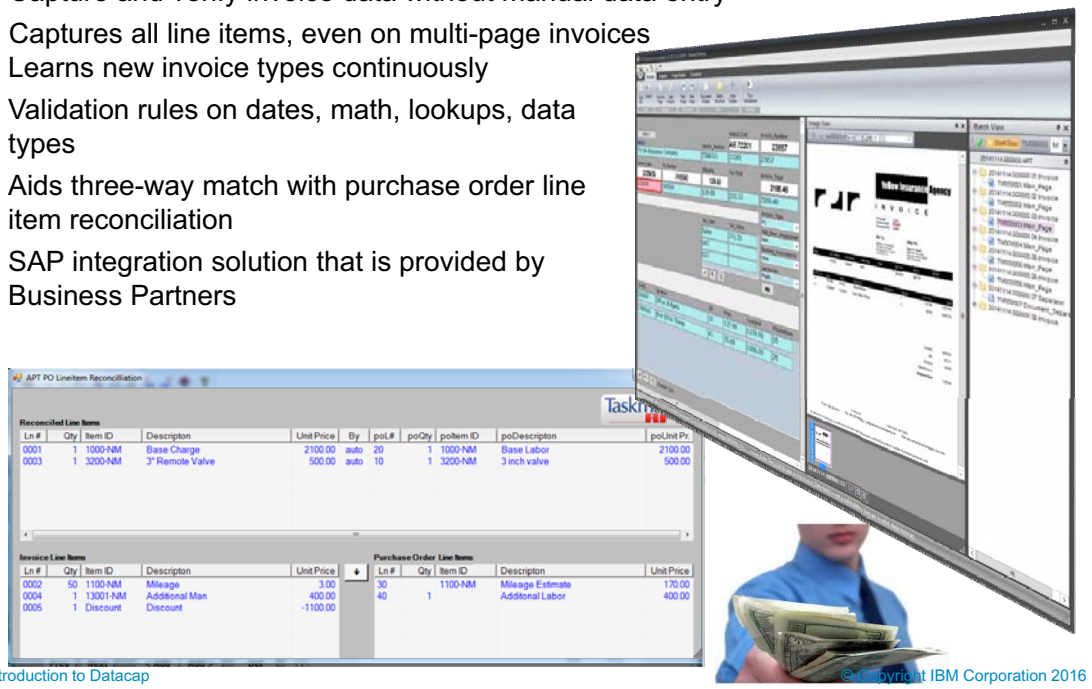
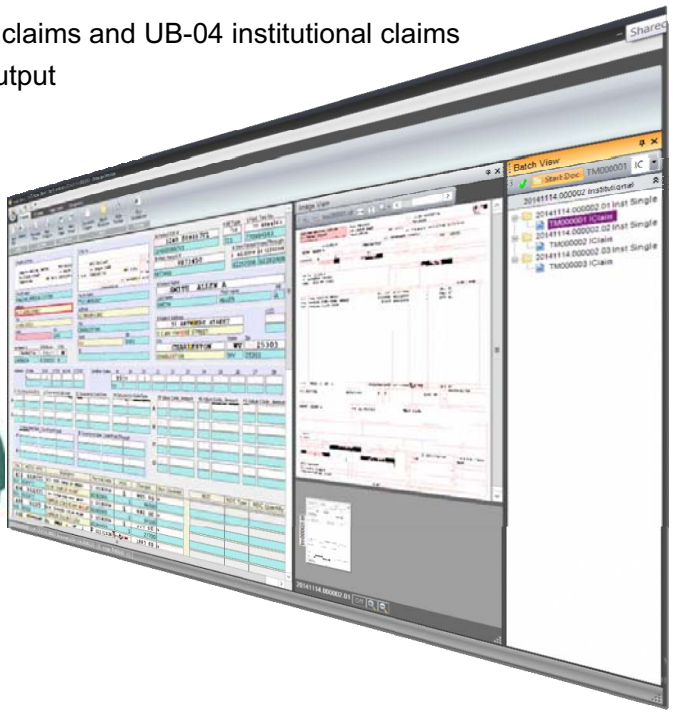


Figure 1-49. Datacap Accounts Payable

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

Figure 1-50. Datacap Medical Claims

IBM Training

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

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.

IBM Training

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

IBM Training

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

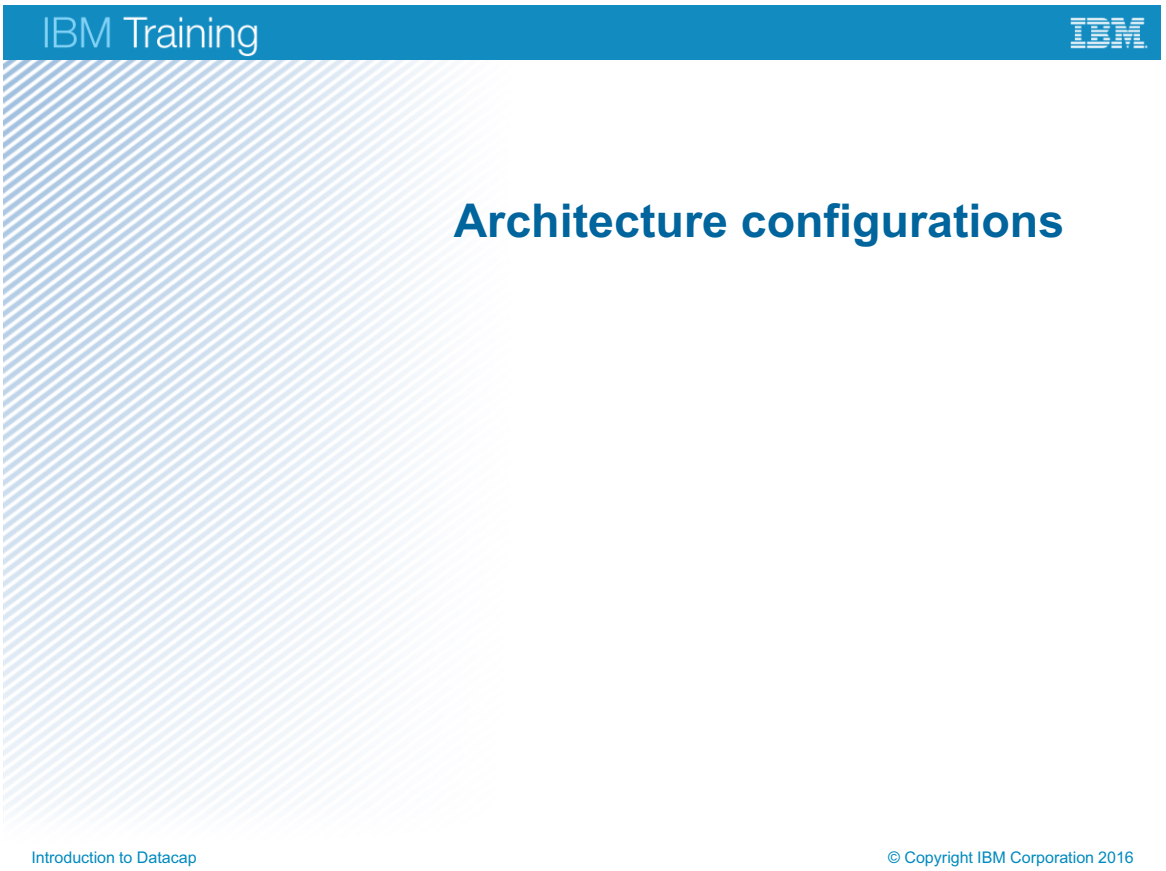


Figure 1-54. Architecture configurations

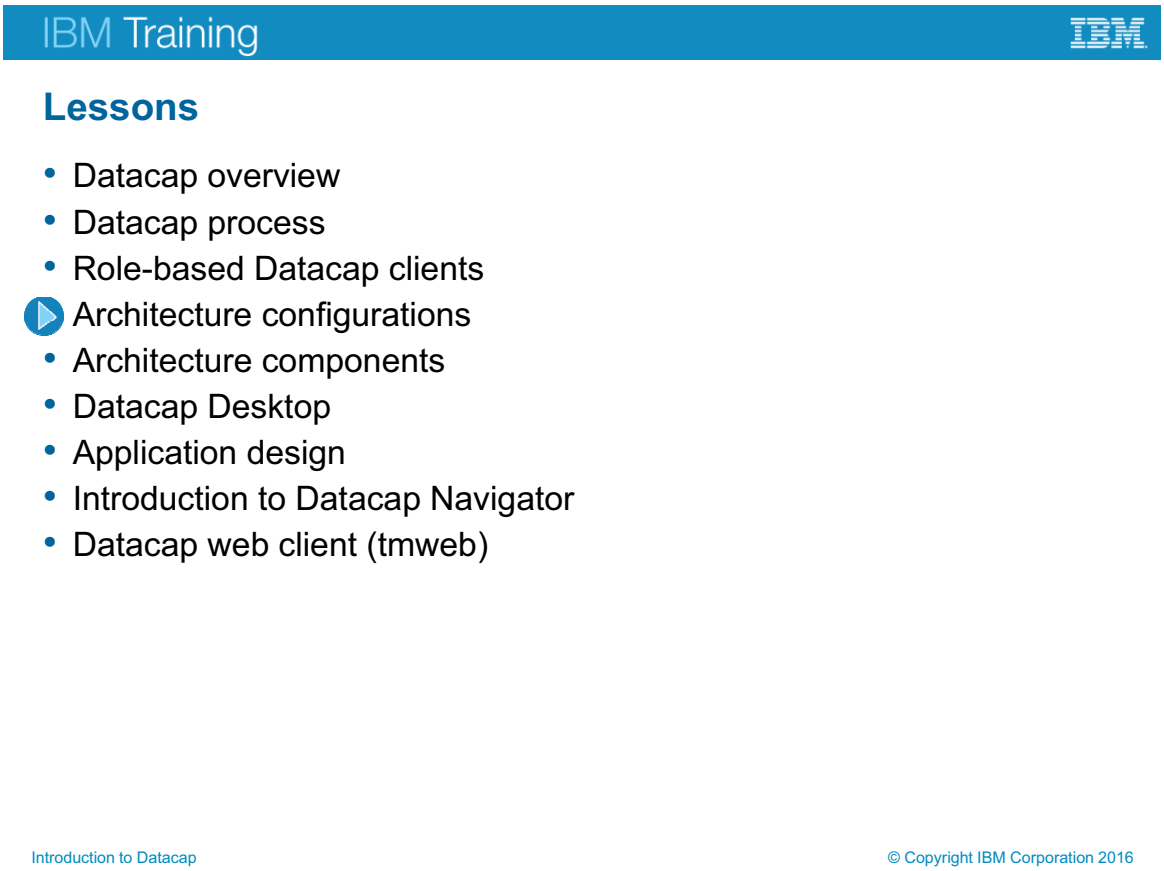


Figure 1-55. Lessons

IBM Training

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?

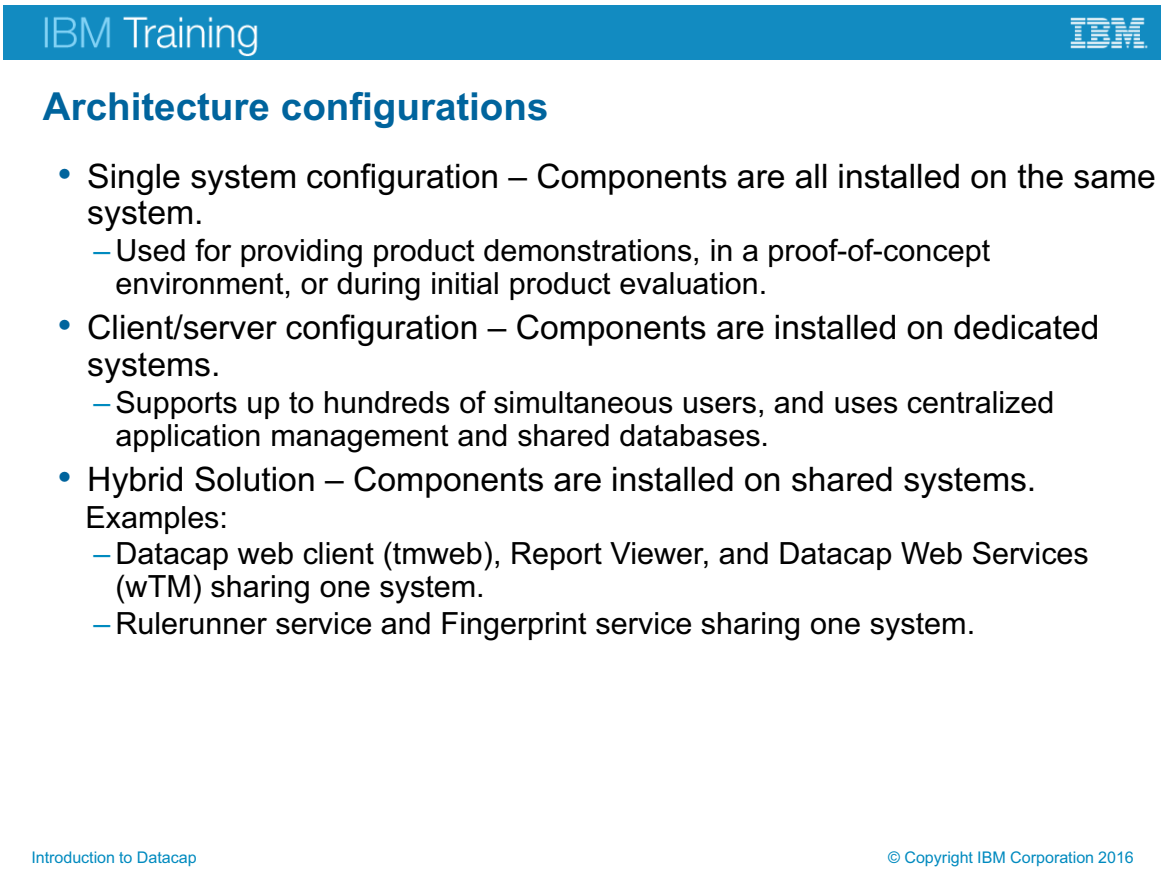


Figure 1-57. Architecture configurations

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.

IBM Training

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.

Introduction to Datacap

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.

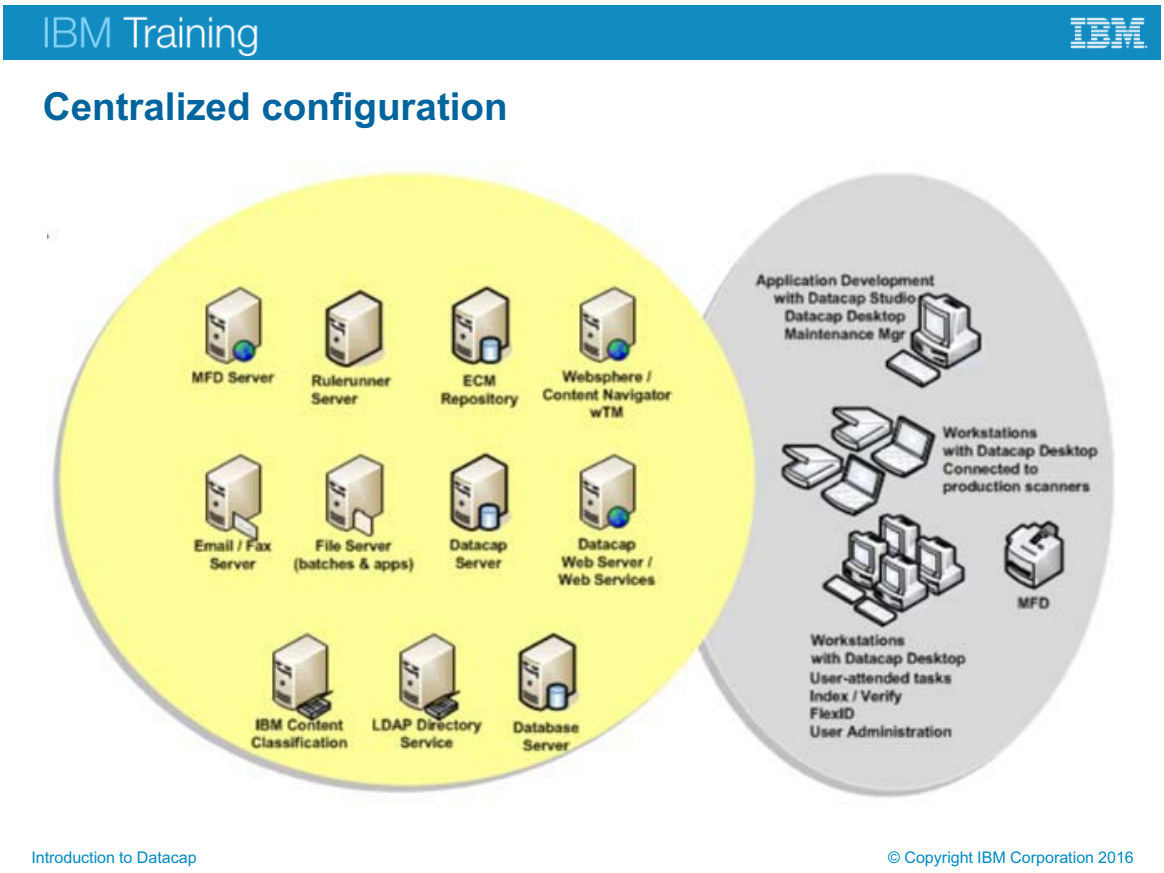


Figure 1-59. Centralized configuration

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.

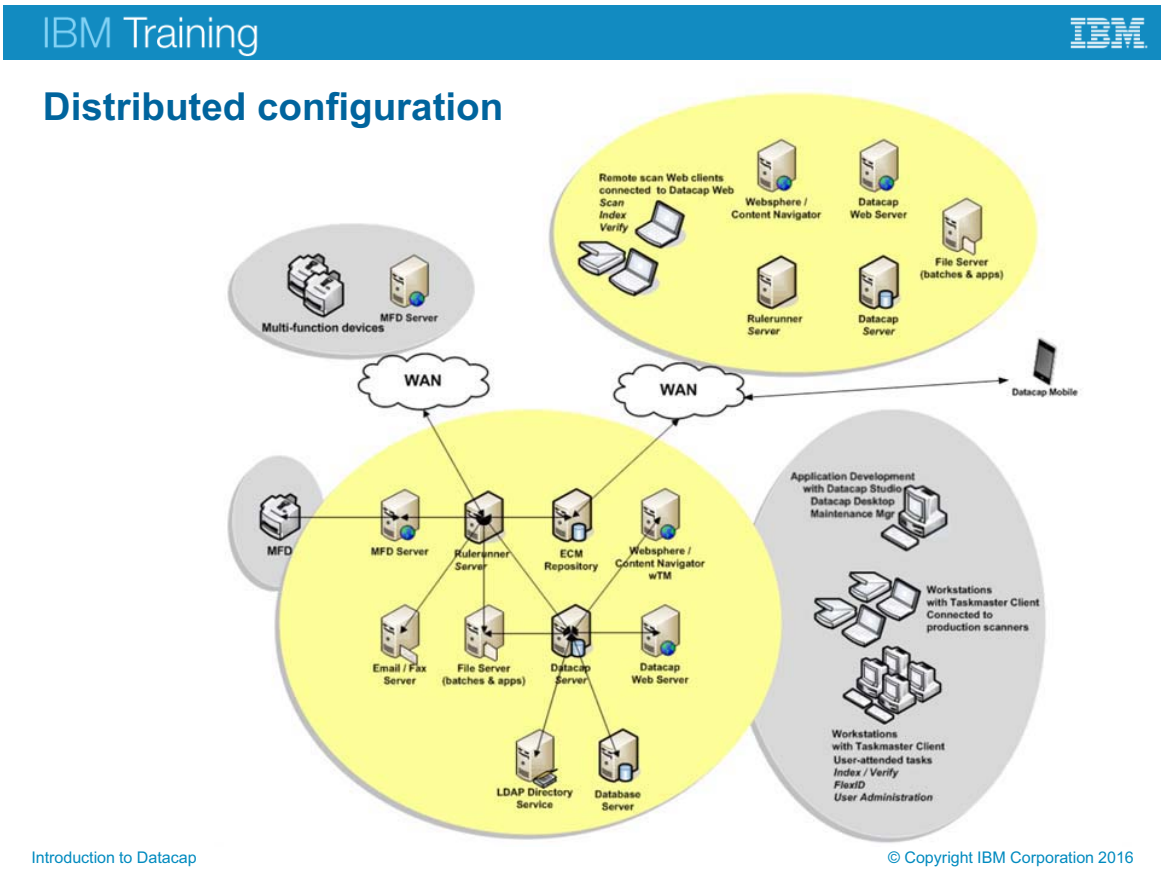


Figure 1-60. Distributed configuration

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.

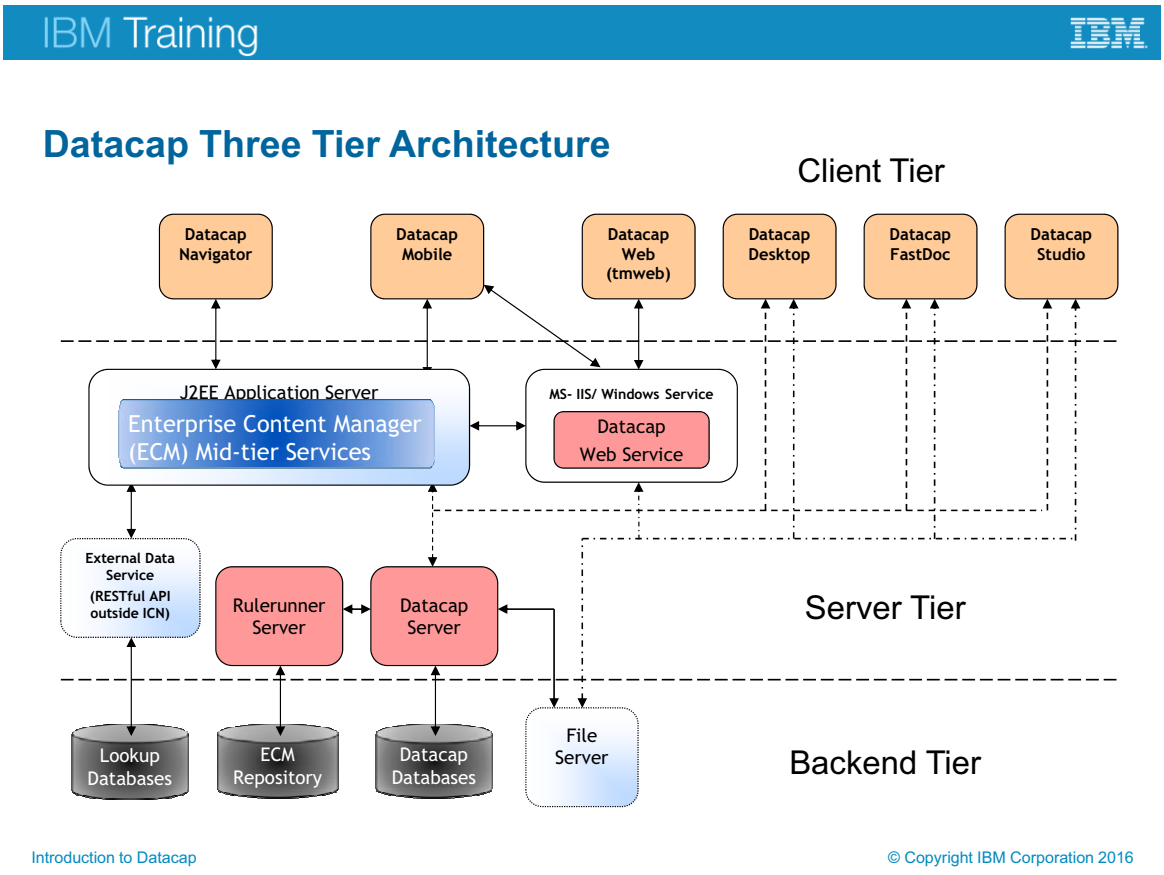


Figure 1-61. Datacap Three Tier Architecture

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

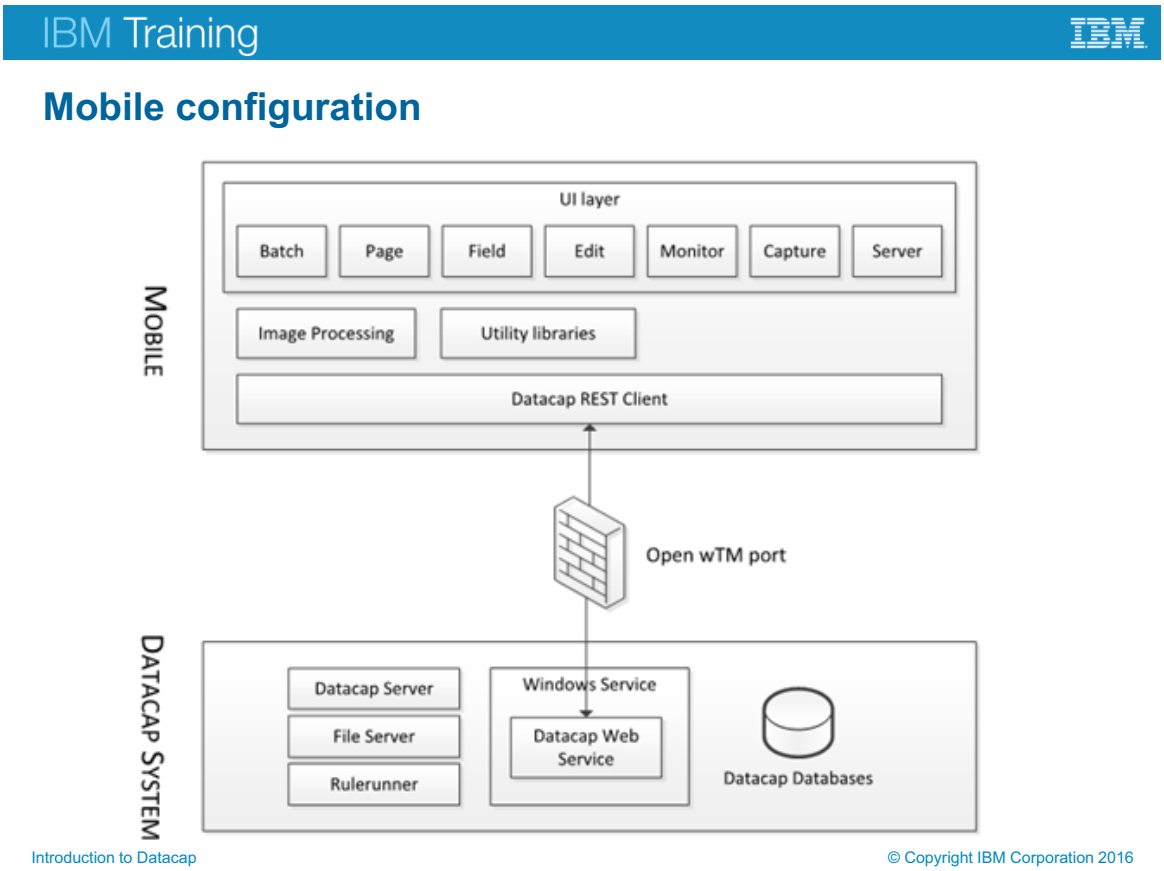


Figure 1-62. Mobile configuration

Datacap mobile architecture with remote access

This illustration shown 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

IBM Training

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)

IBM Training

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)

IBM Training

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)

IBM Training

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)

IBM Training

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

Figure 1-67. Review questions (3)

IBM Training

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

Figure 1-68. Review answers (3)

Lesson 1.5. Architecture components

Figure 1-69. Architecture components

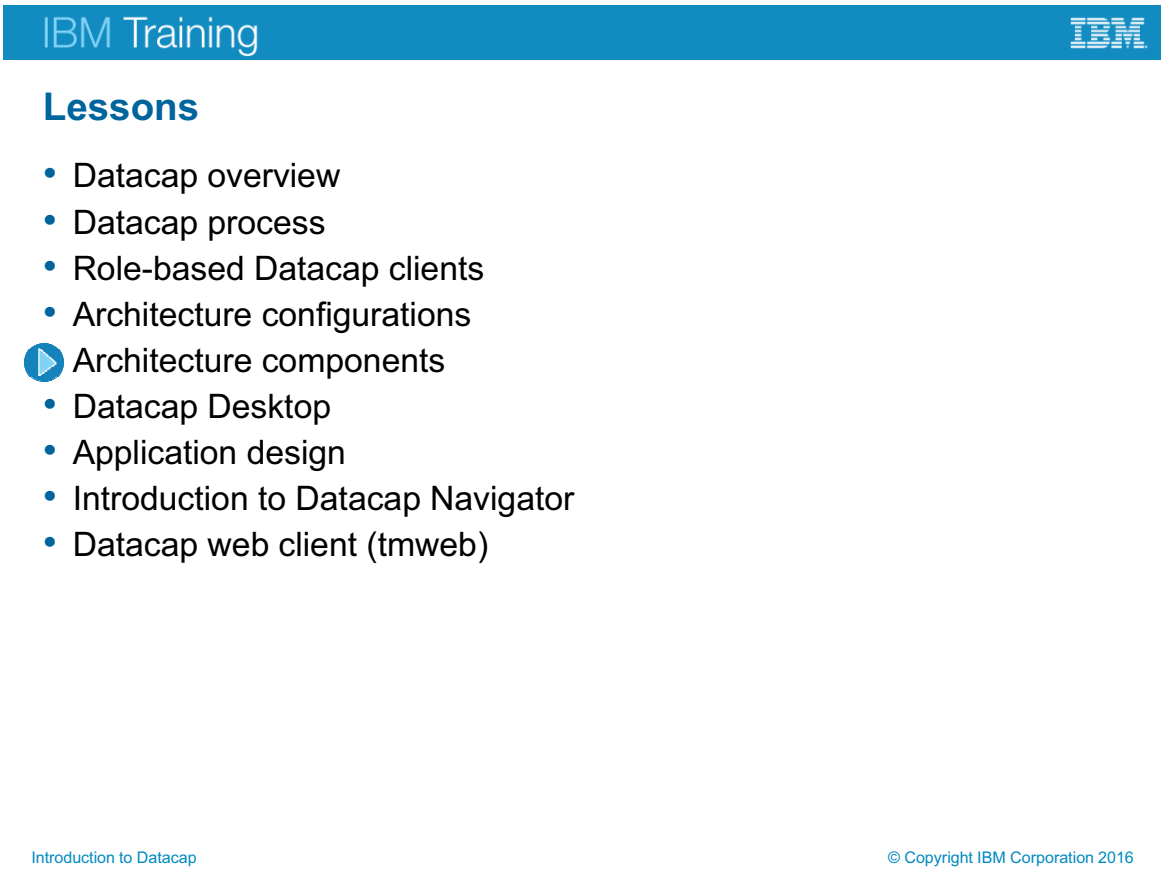


Figure 1-70. Lessons

IBM Training

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?

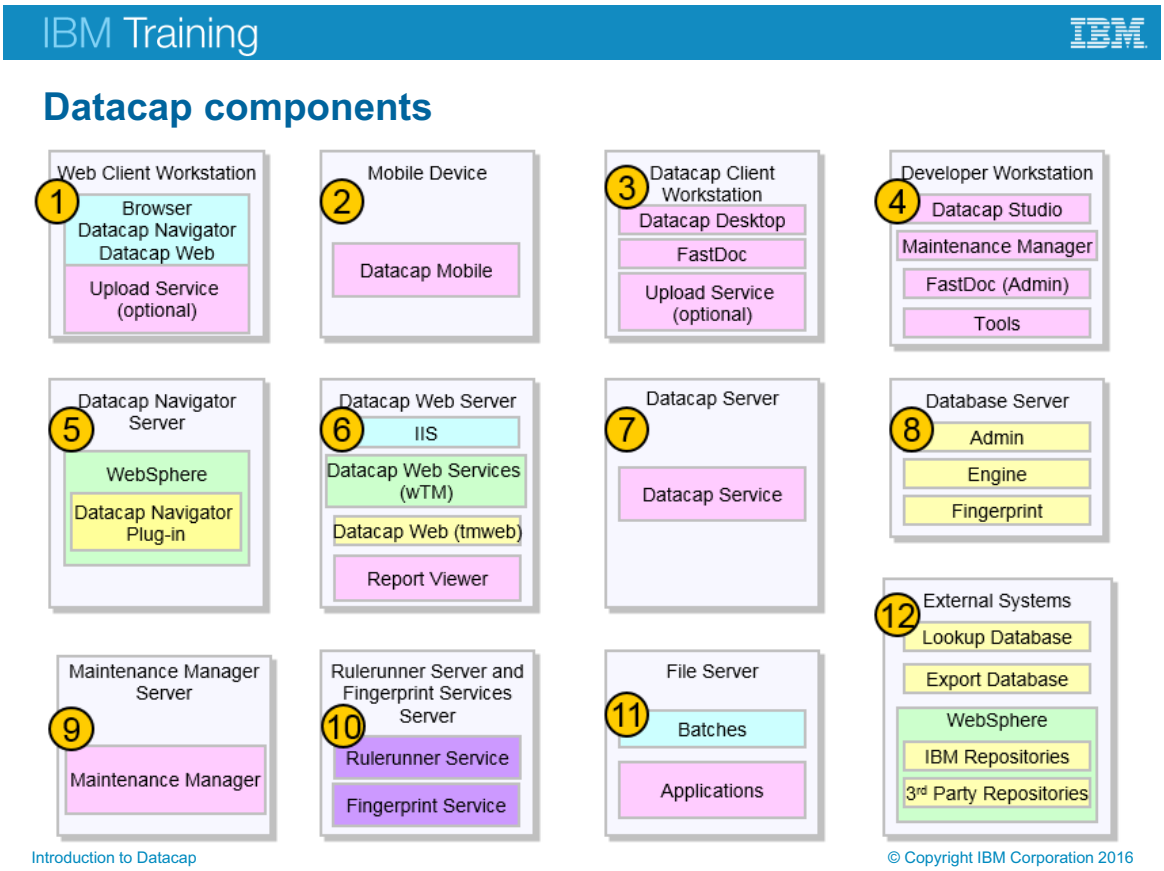


Figure 1-72. Datacap components

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

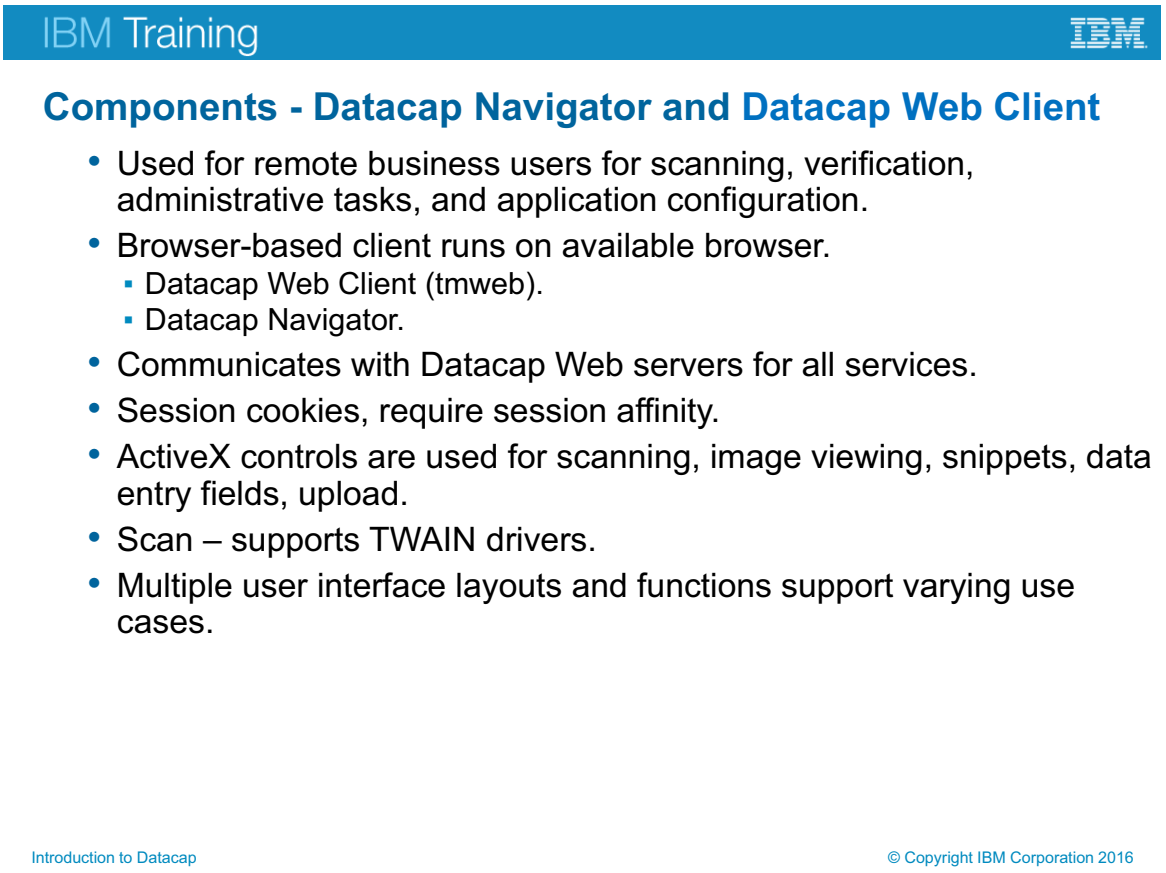


Figure 1-73. Components - Datacap Navigator and Datacap Web Client

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.

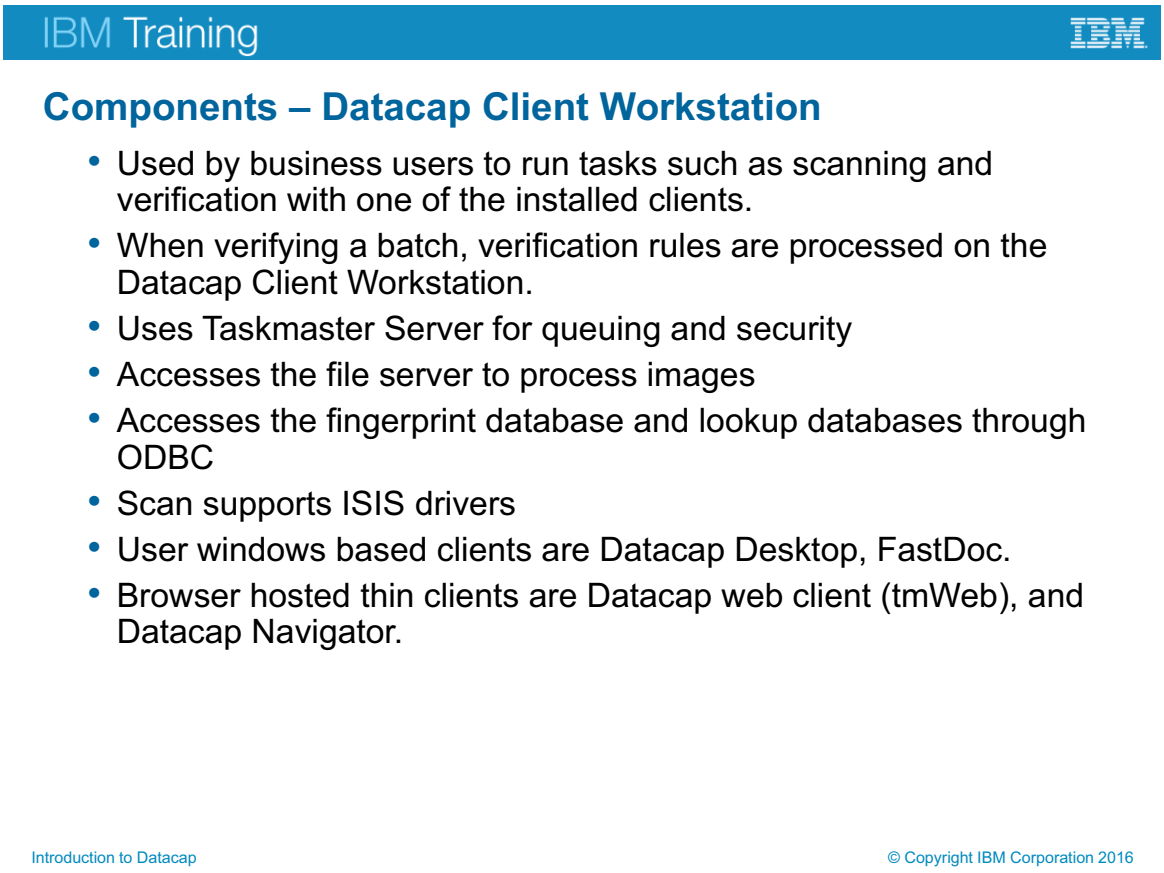


Figure 1-74. Components – Datacap Client Workstation

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.

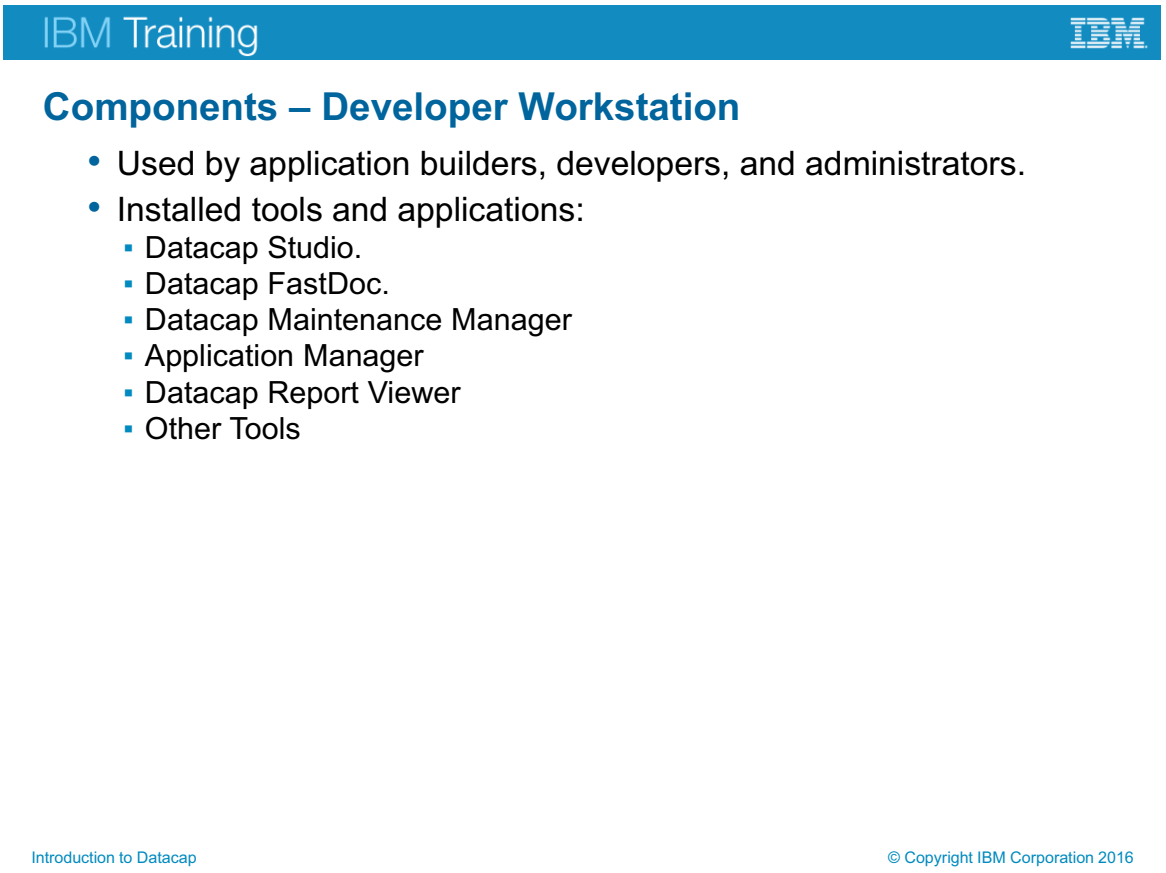


Figure 1-75. Components – Developer Workstation

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.

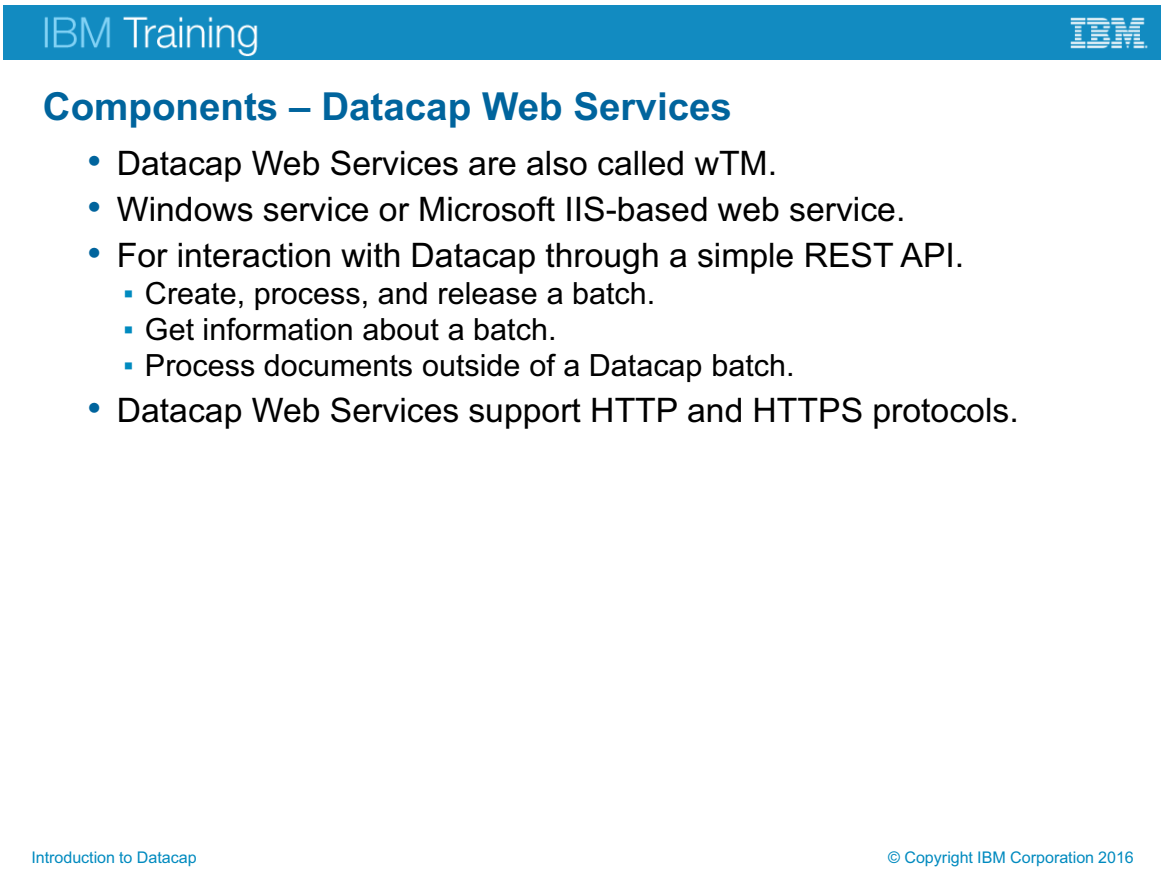


Figure 1-78. Components – Datacap Web Services

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

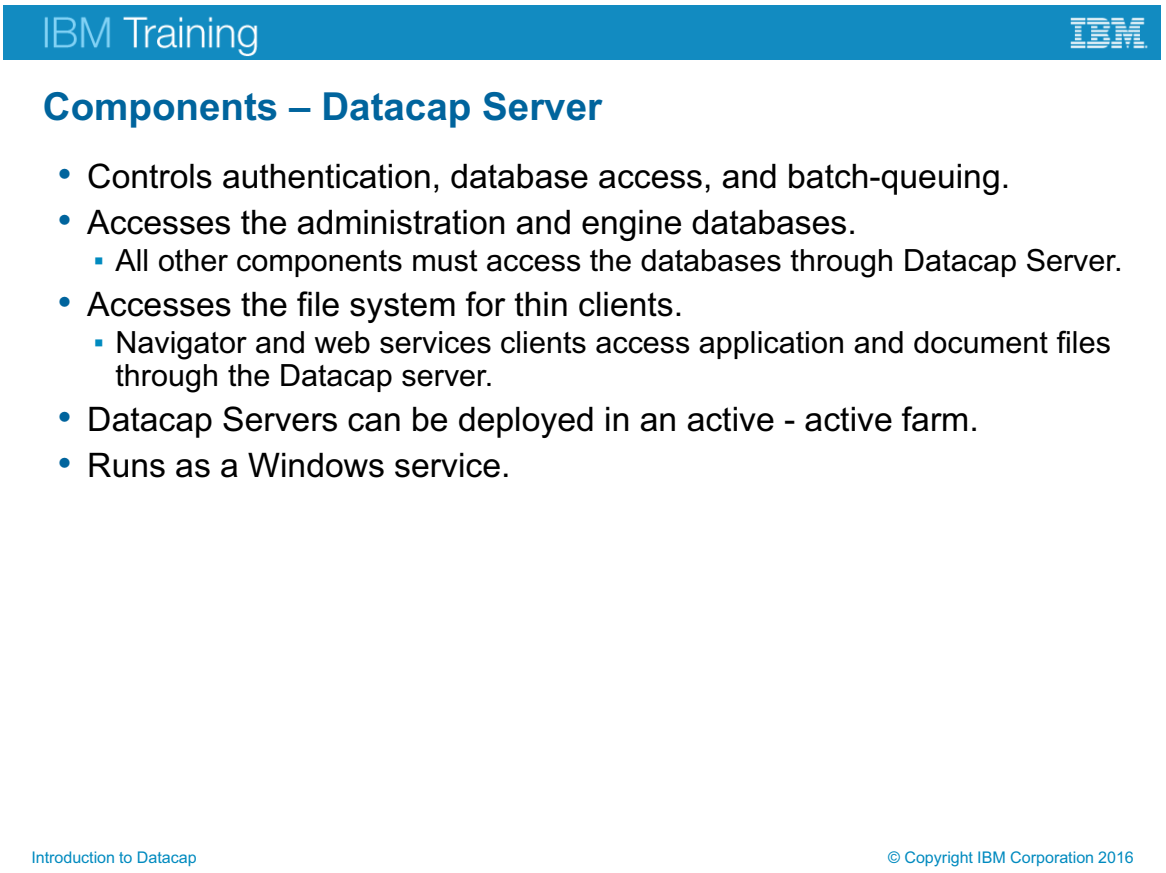


Figure 1-79. Components – Datacap Server

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.

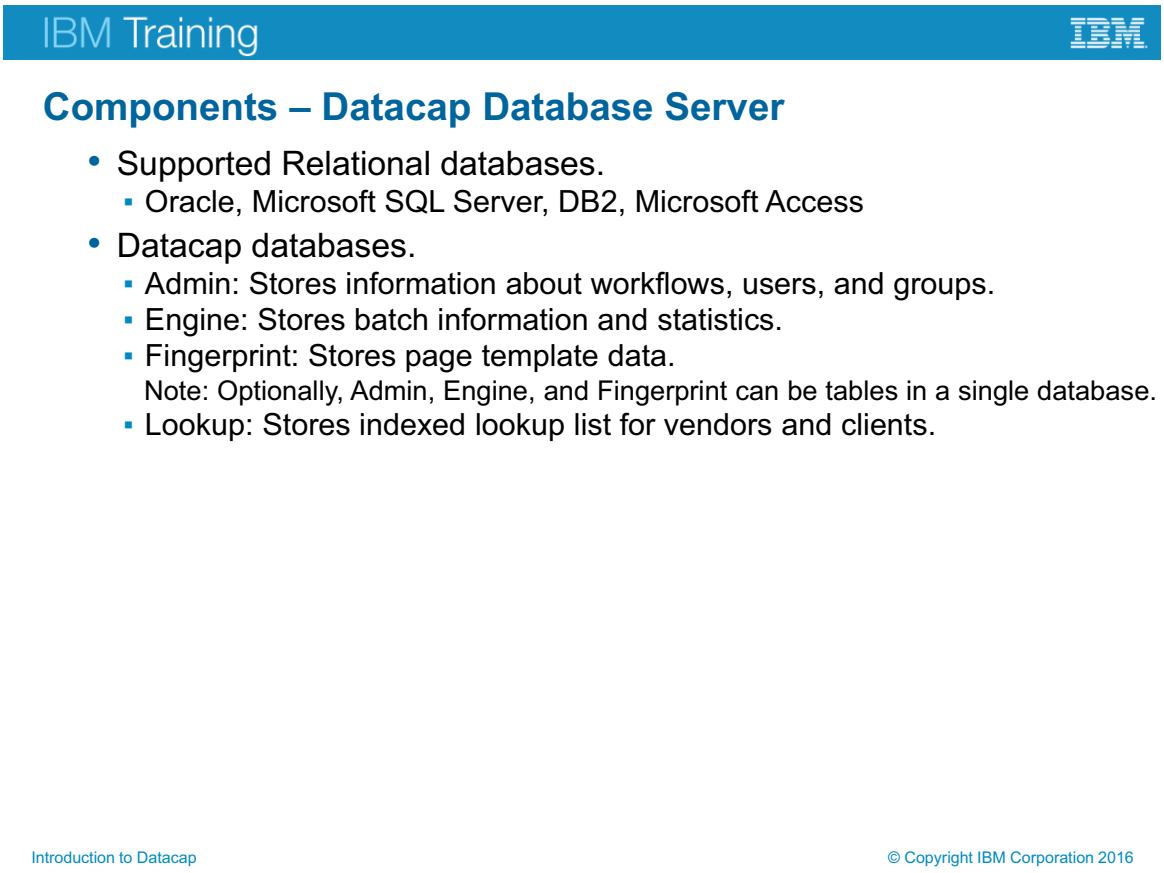


Figure 1-80. Components – Datacap Database Server

Database Server

- One ore 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.

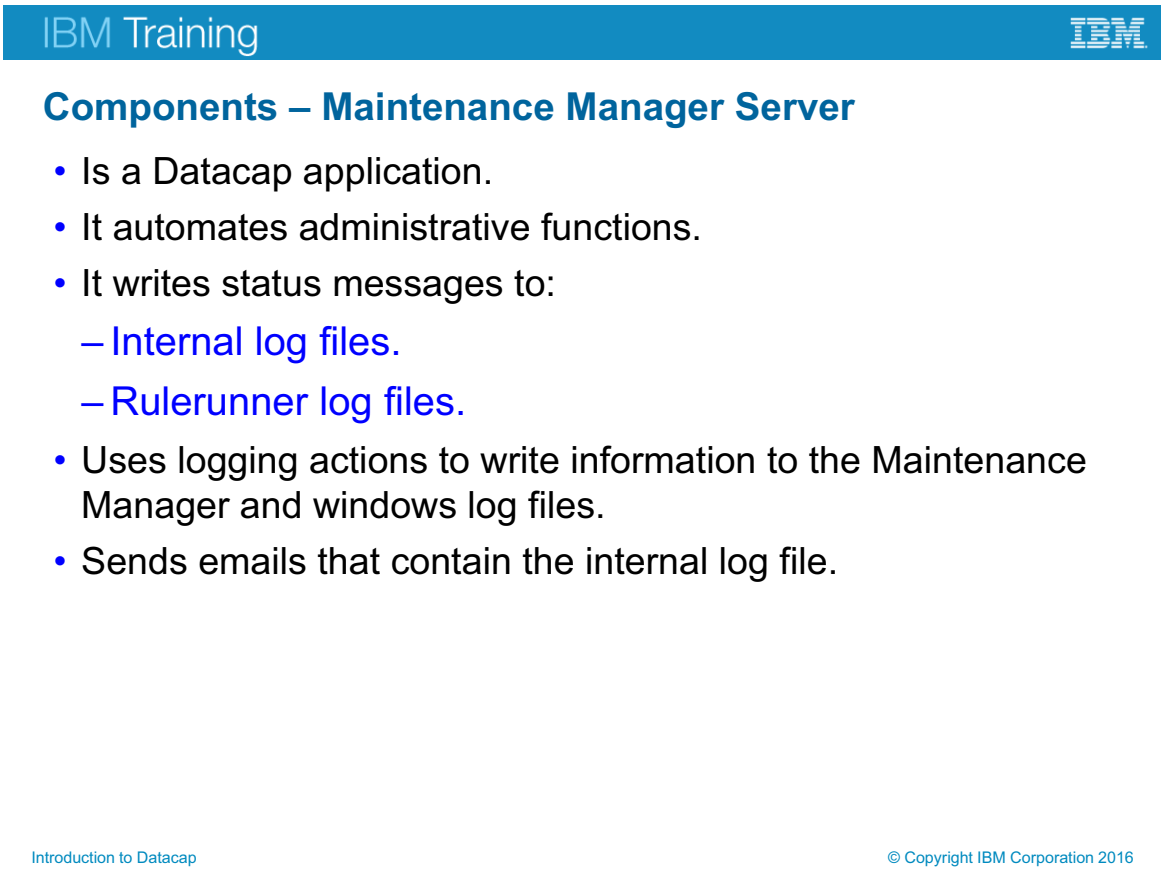


Figure 1-81. Components – Maintenance Manager Server

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.

IBM Training

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

Introduction to Datacap

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.

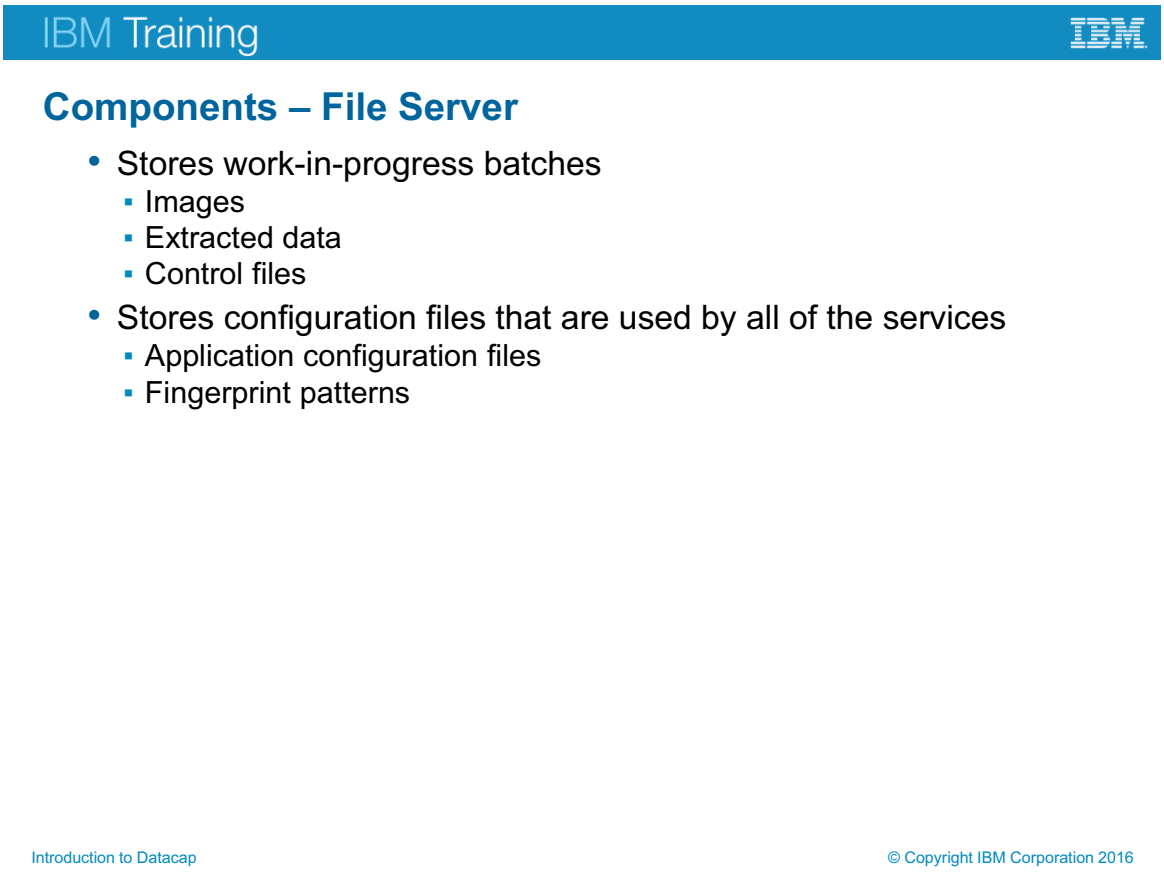


Figure 1-83. Components – File Server

File server

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

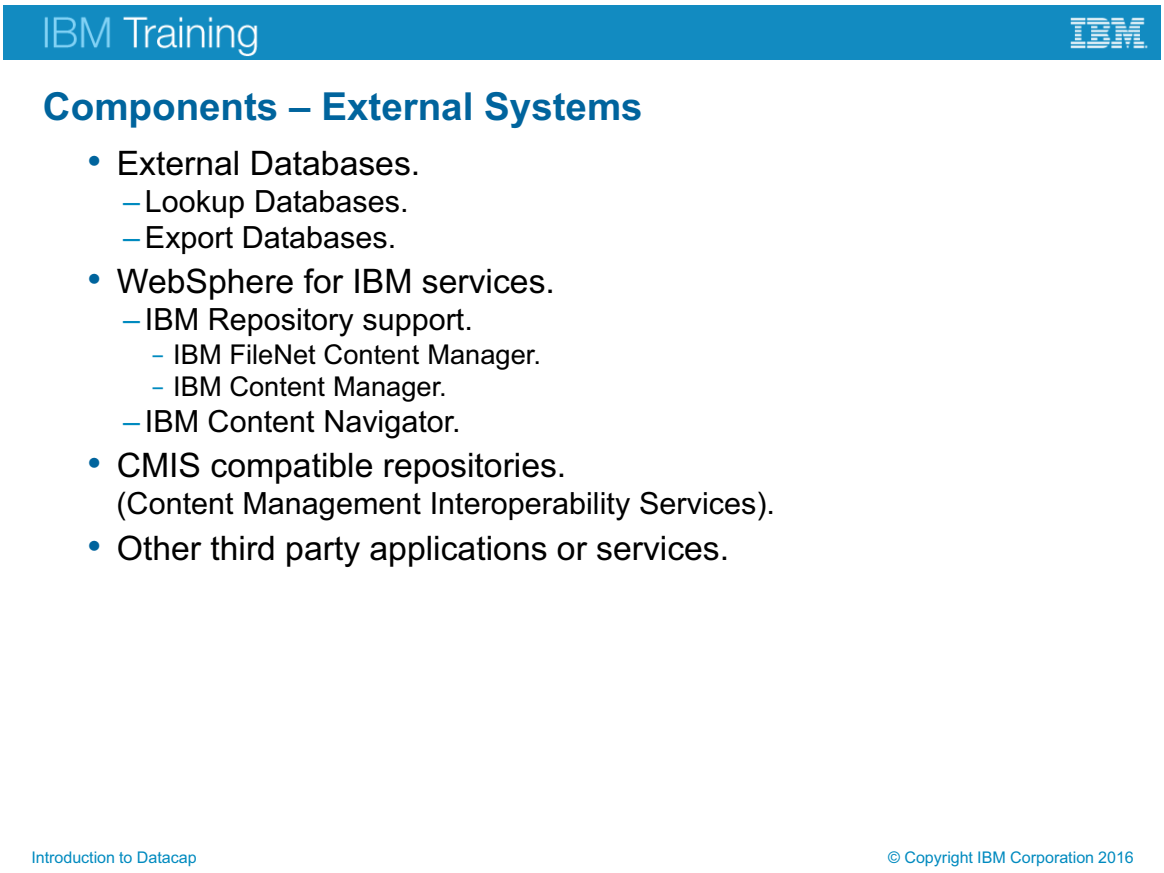


Figure 1-84. Components – External Systems

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 provid connectivity to external system.

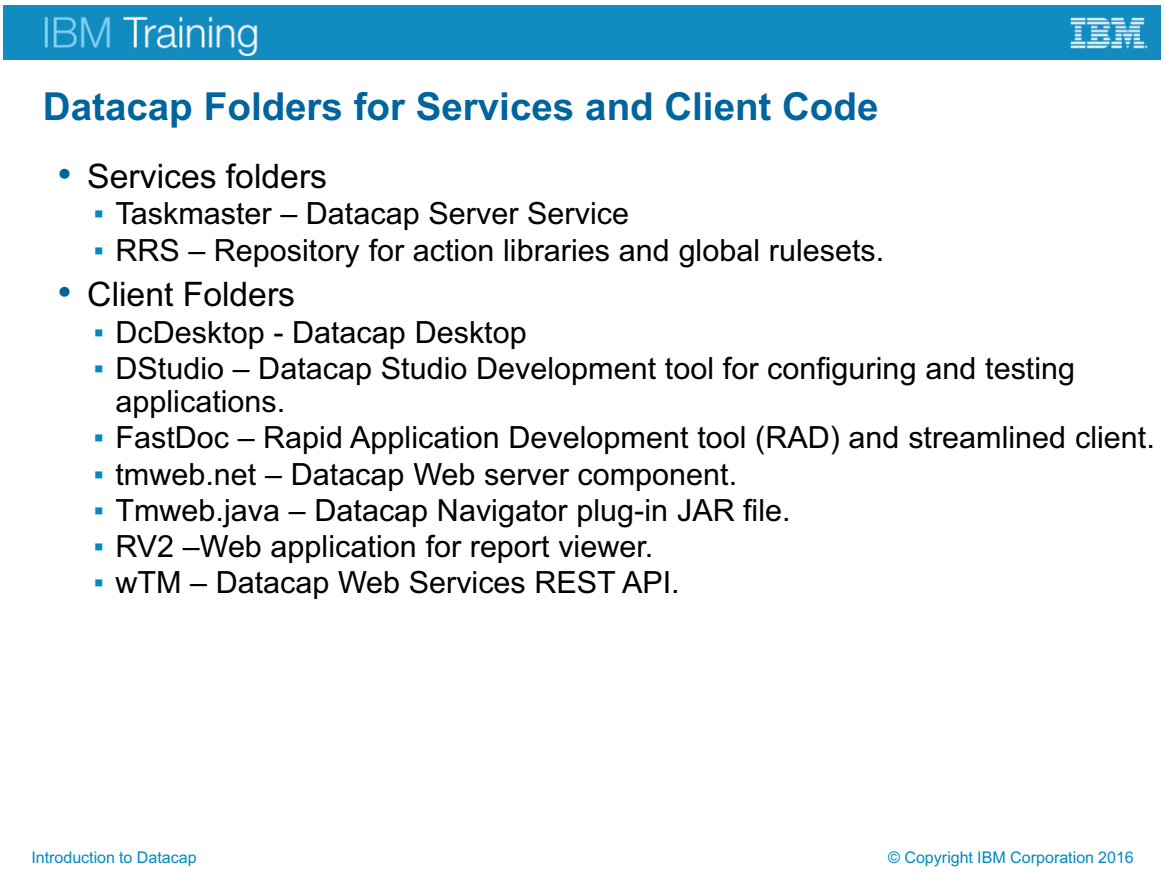


Figure 1-85. Datacap Folders for Services and Client Code

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

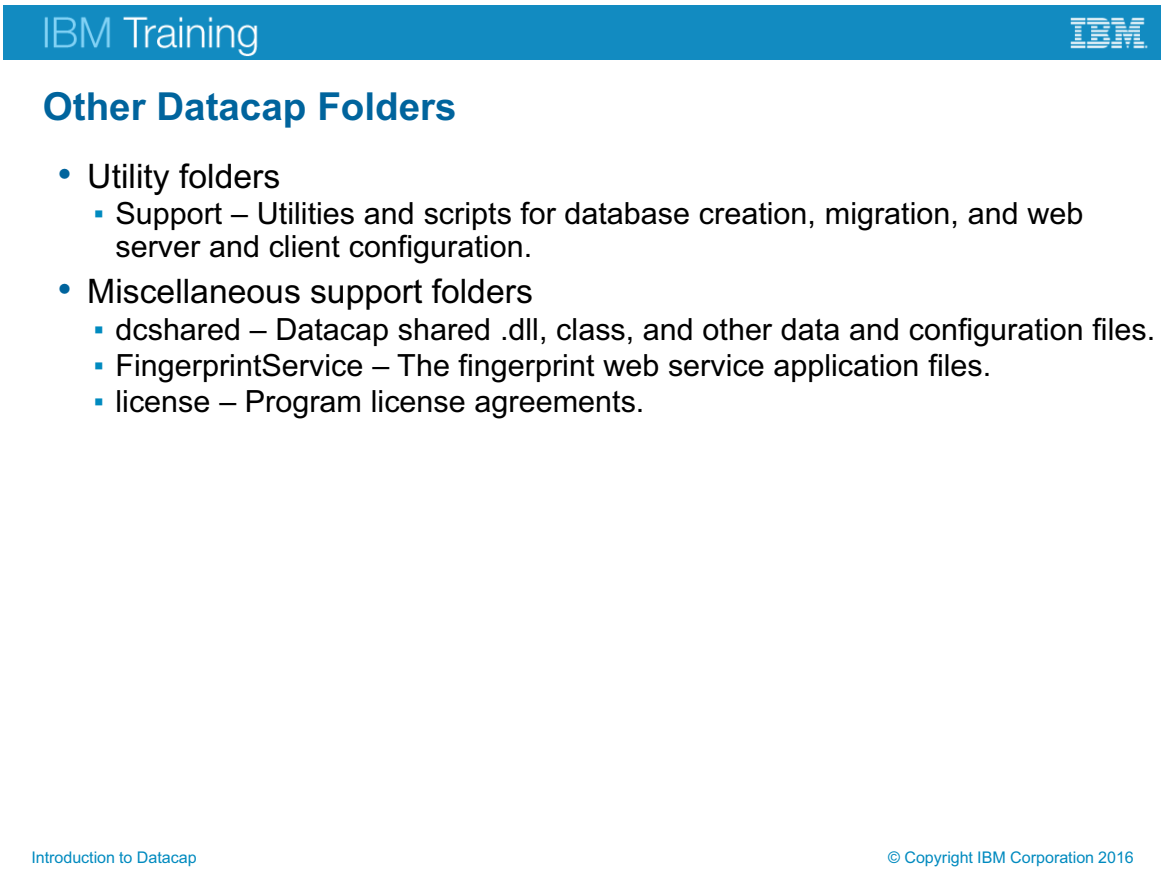


Figure 1-86. Other Datacap Folders

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

IBM Training

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.

datacap.xml

```
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

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

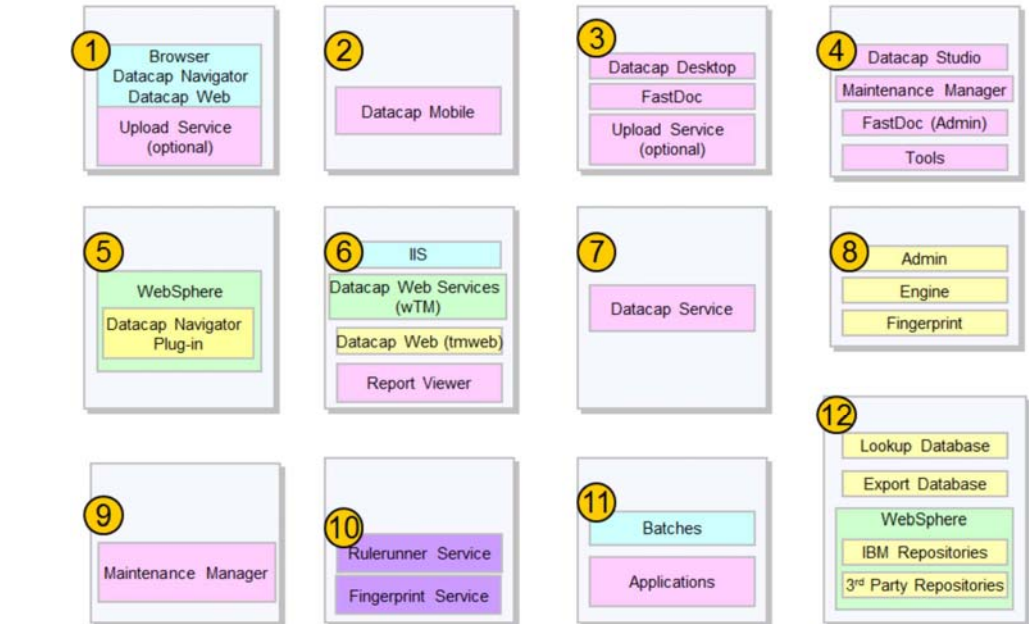
- The datacap.xml file is in the Datacap root folder. It contains the Datacap version number and a tag that defines the name for each of the defined Datacap Capture applications.
- The version and exact build number are also shown in the C:\Datacap\version.txt file.

  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.

IBM Training

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.



Introduction to Datacap

Figure 1-88. Review question (1)

Write your answers here:

Datacap component nameComponent 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

IBM Training

Review answers (1)

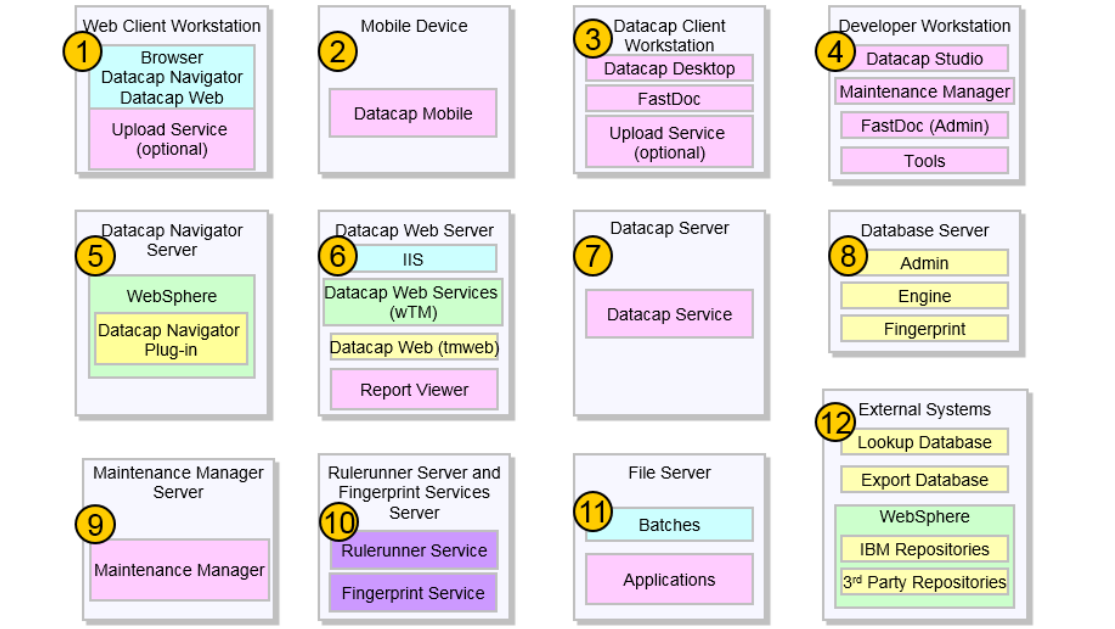


Figure 1-89. Review answers (1)

The answers are:

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

Lesson 1.6. Datacap Desktop

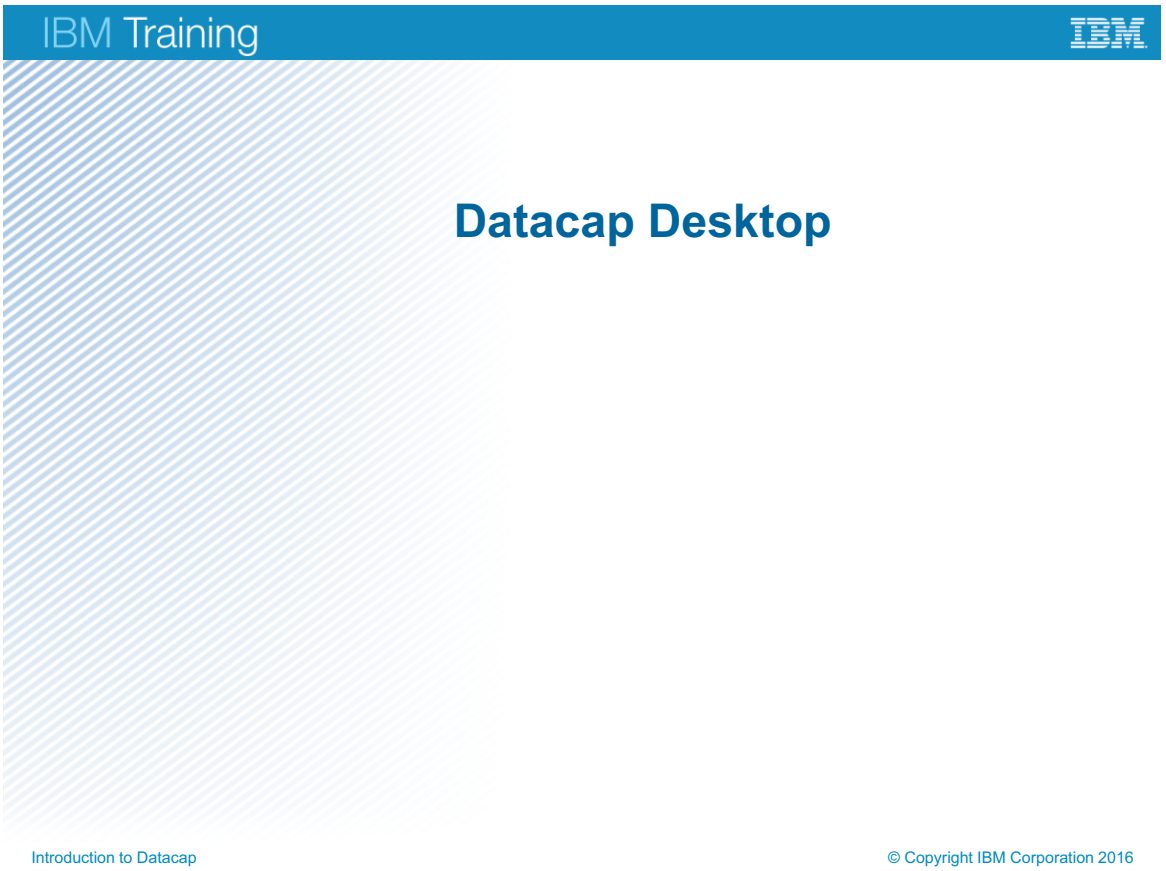


Figure 1-90. Datacap Desktop

Figure 1-91. Lessons

IBM Training

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?

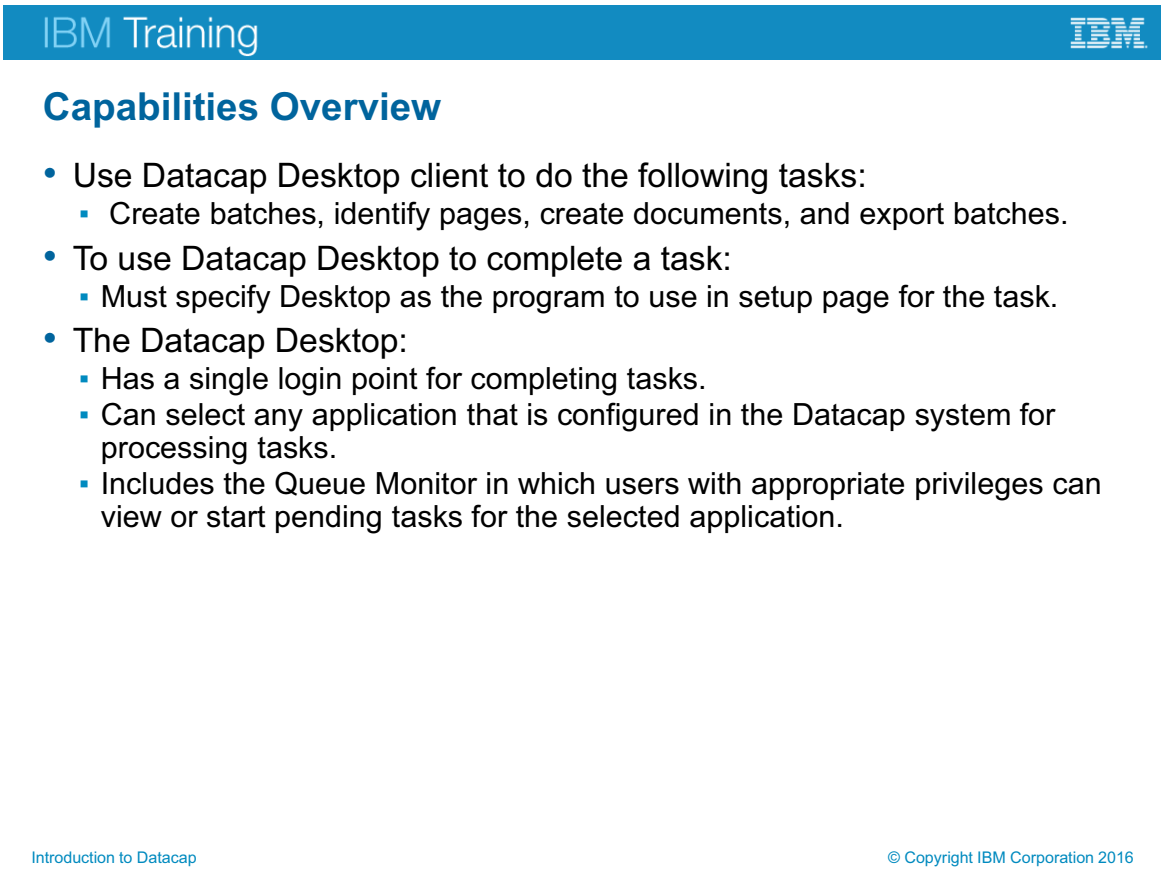


Figure 1-93. Capabilities Overview

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 categorizes for ease of access. The next step might also be direct them to subsequent application for further processing.

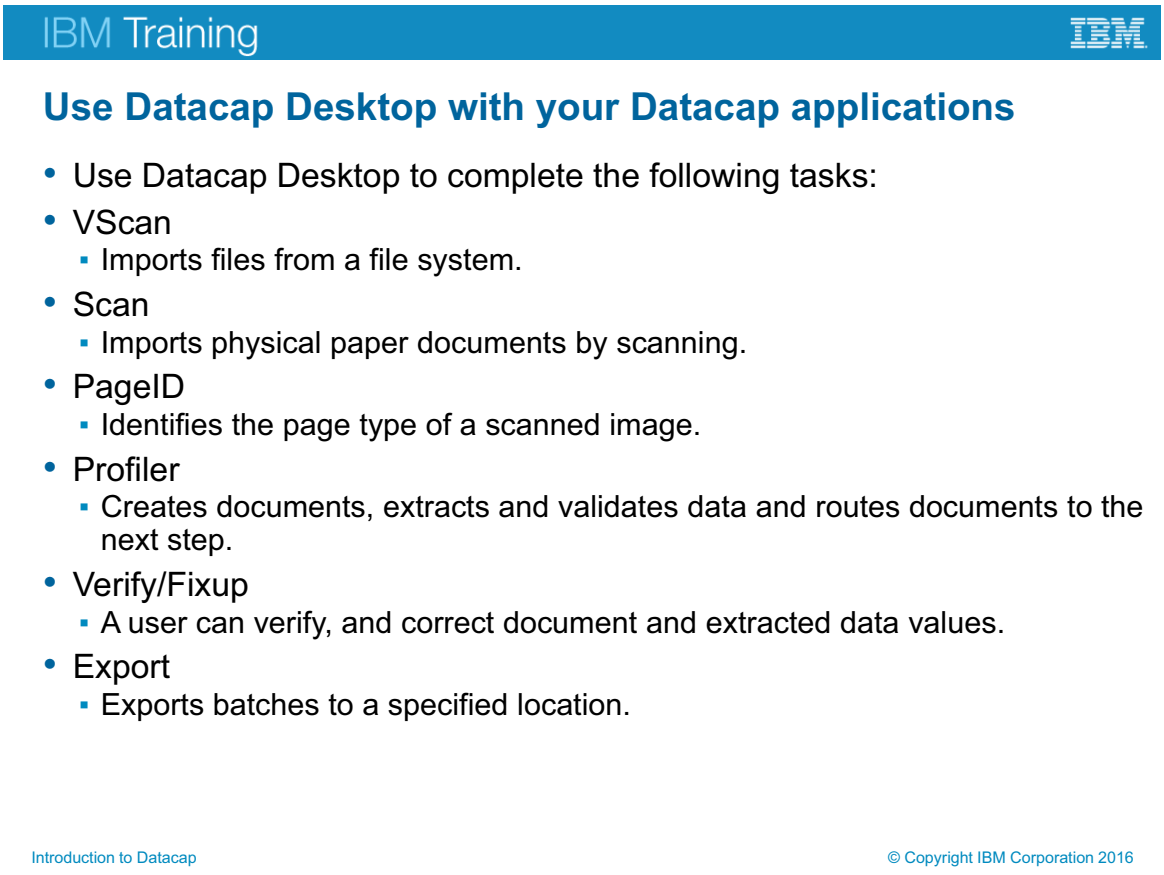


Figure 1-94. Use Datacap Desktop with your Datacap applications

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.

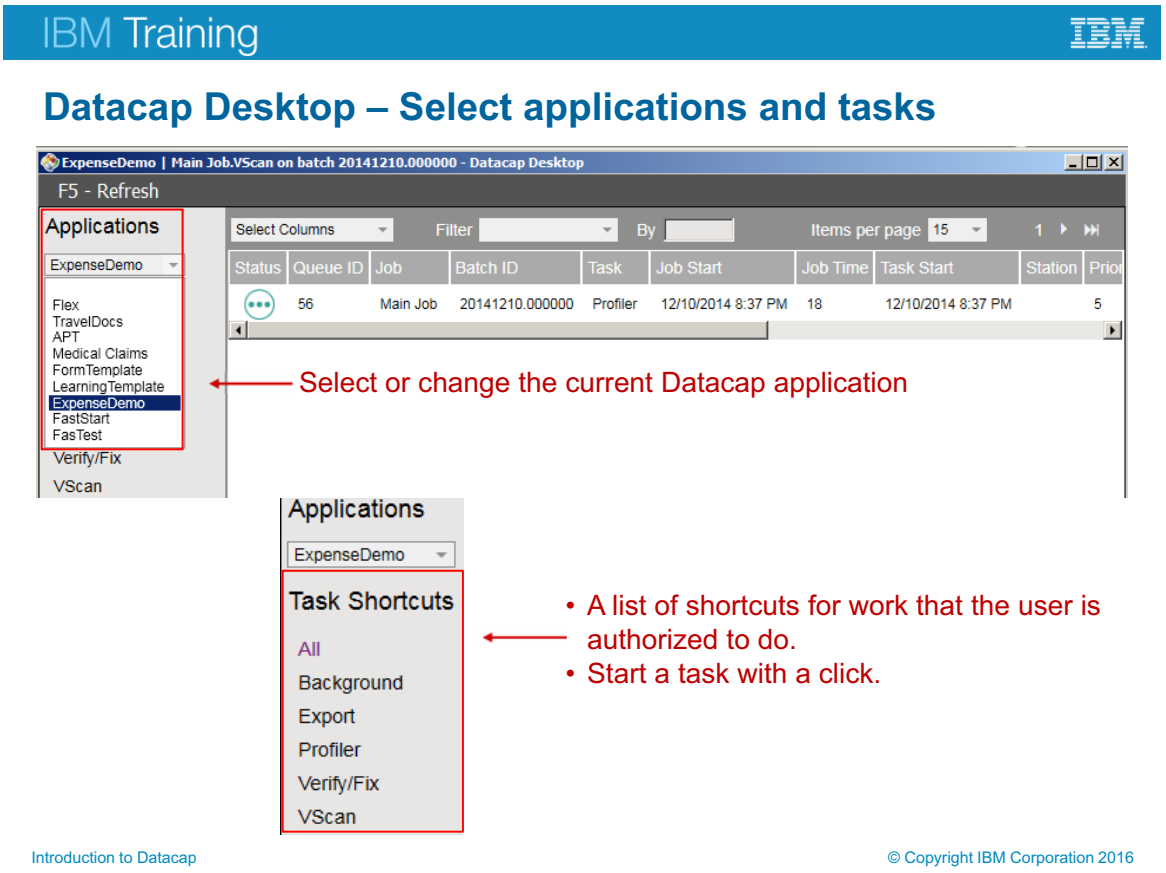


Figure 1-95. Datacap Desktop – Select applications and tasks

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

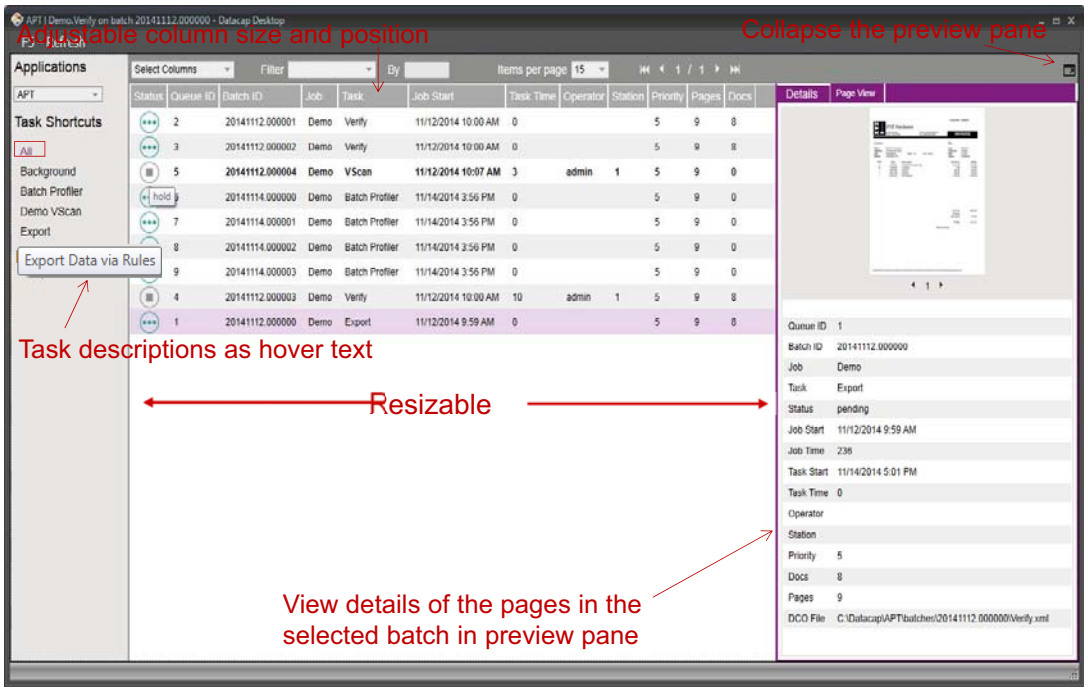
- Users can select an application or change to another application from the list.
- After selecting a particular application, a list of allowed shortcuts for work that the user is authorized to do is displayed below the application.
- When the user clicks a shortcut to start, what gets displayed in the 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.

IBM Training

Datacap Desktop - Monitor view



Introduction to Datacap

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.

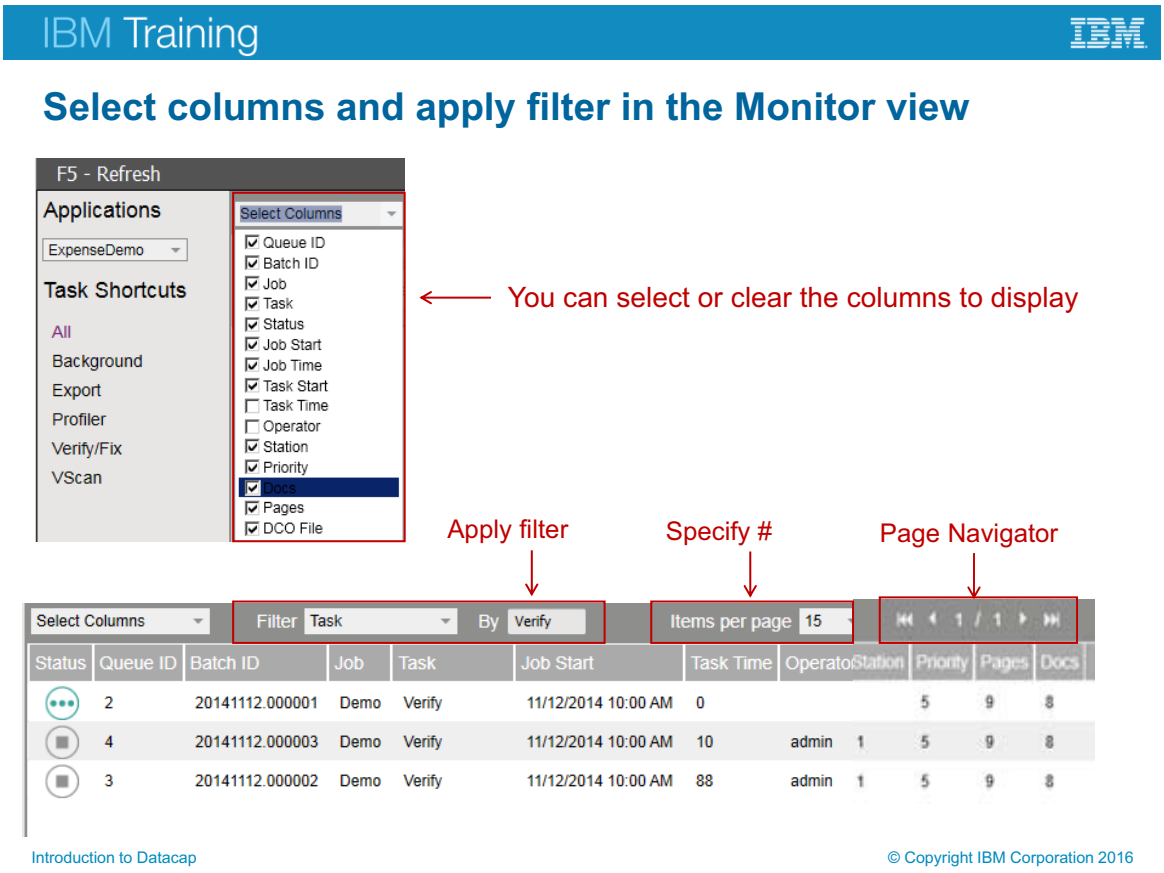


Figure 1-97. Select columns and apply filter in the Monitor view

- 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.

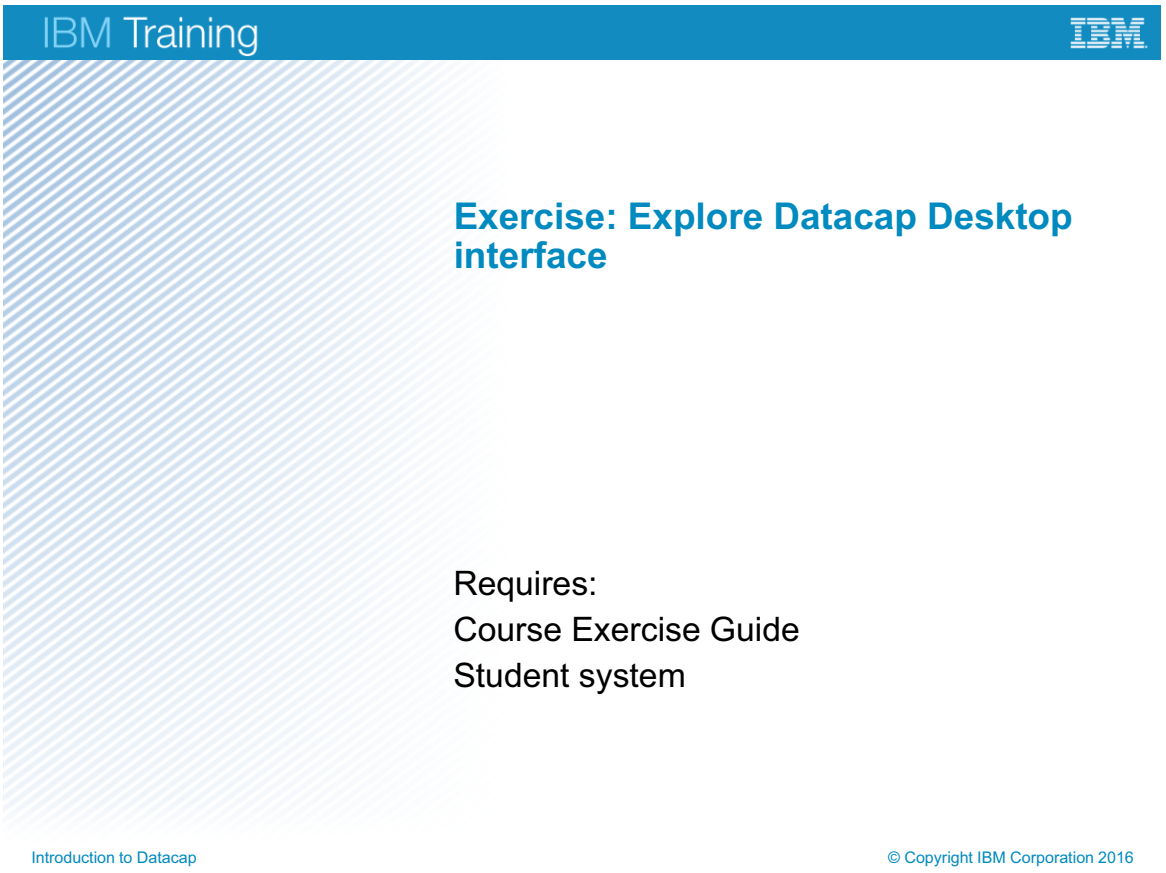


Figure 1-98. Exercise: Explore Datacap Desktop interface

IBM Training

Exercise objectives

- Explore Datacap Desktop interface.

Figure 1-99. Exercise objectives

Lesson 1.7. Application design

Figure 1-100. Application design

Figure 1-101. Lessons

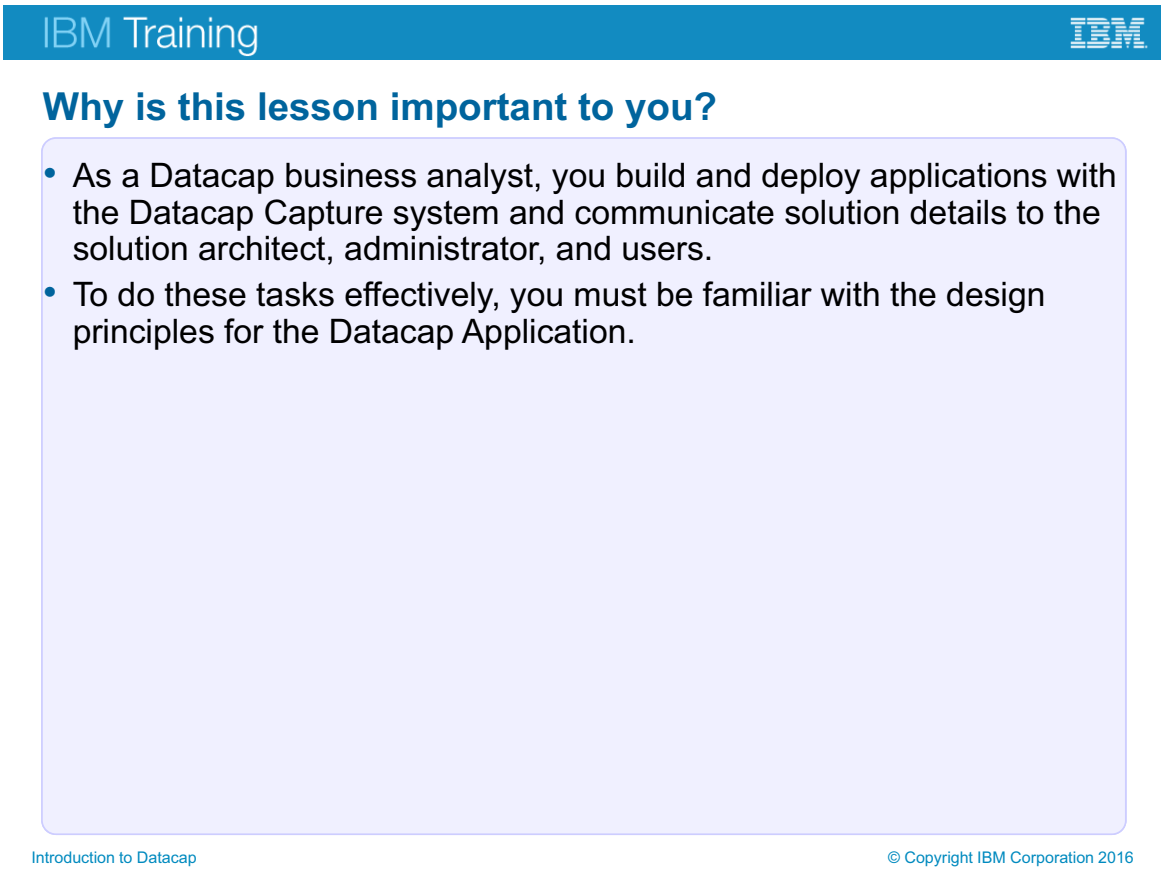


Figure 1-102. Why is this lesson important to you?

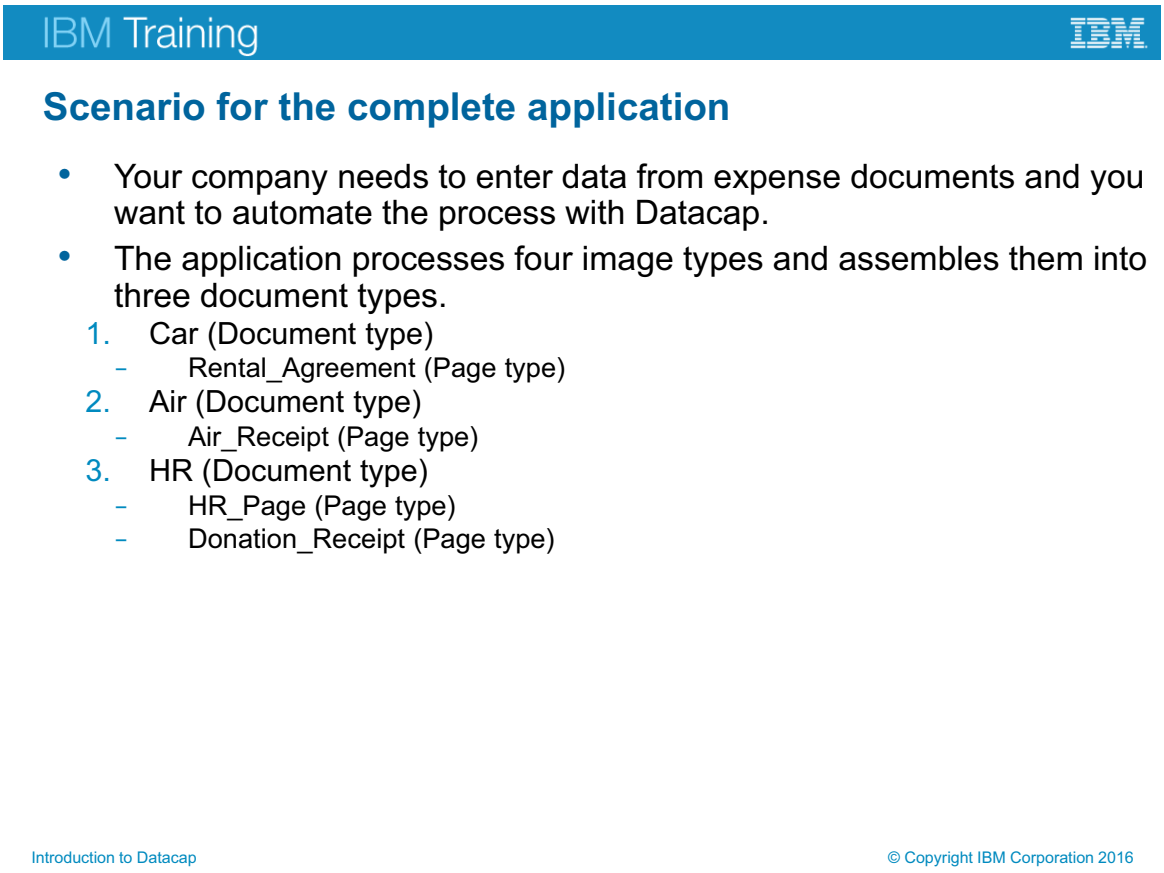


Figure 1-103. Scenario for the complete application

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.

IBM Training

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

IBM Training

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

IBM Training

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

IBM Training

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

IBM Training

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

IBM Training

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

IBM Training

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

Introduction to Datacap

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.

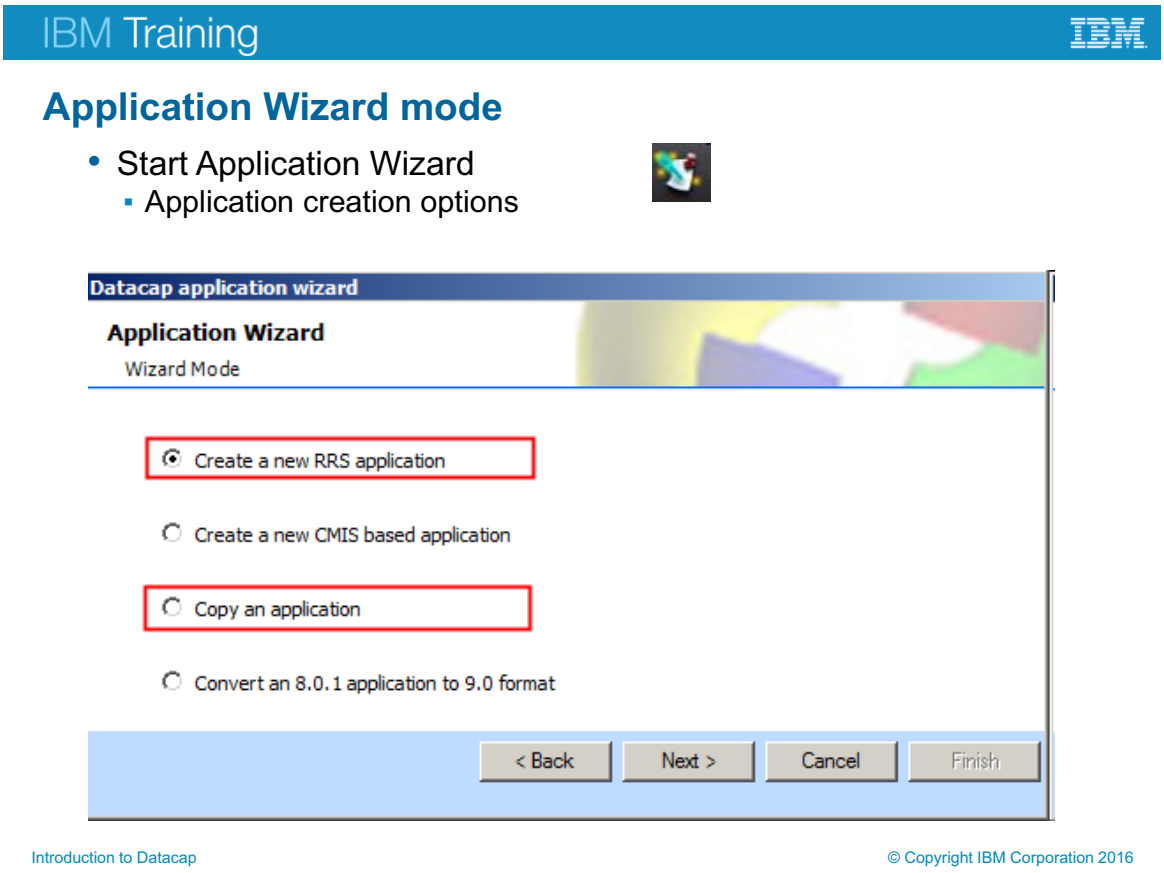


Figure 1-111. Application Wizard mode

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.

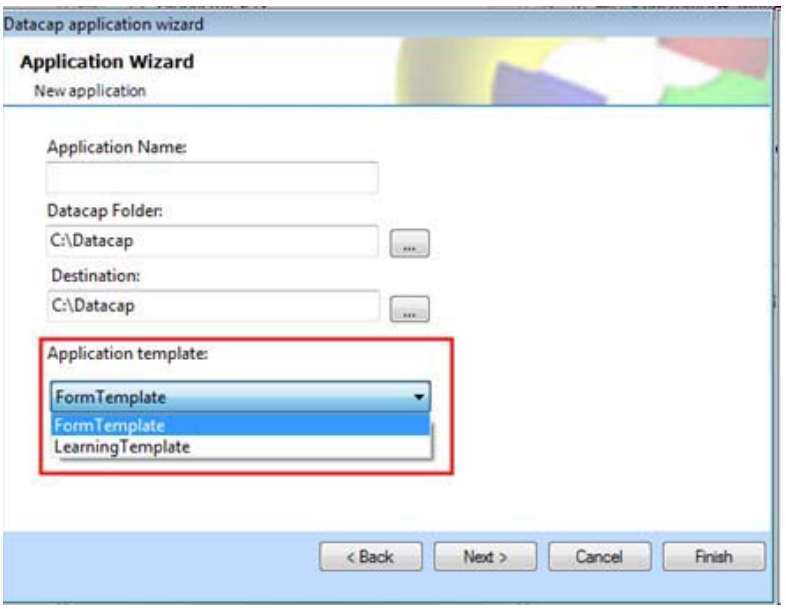
- Create an RRS application
  - Use this option to create a new allocation skeleton from one of the built-in application templates.
- Create a CMIS-based application
  - Use this option to create a new allocation skeleton from one of the built-in application templates.
  - Configured to use CMIS (Common 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.

IBM Training

Application Wizard New Application window

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



Introduction to Datacap

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.

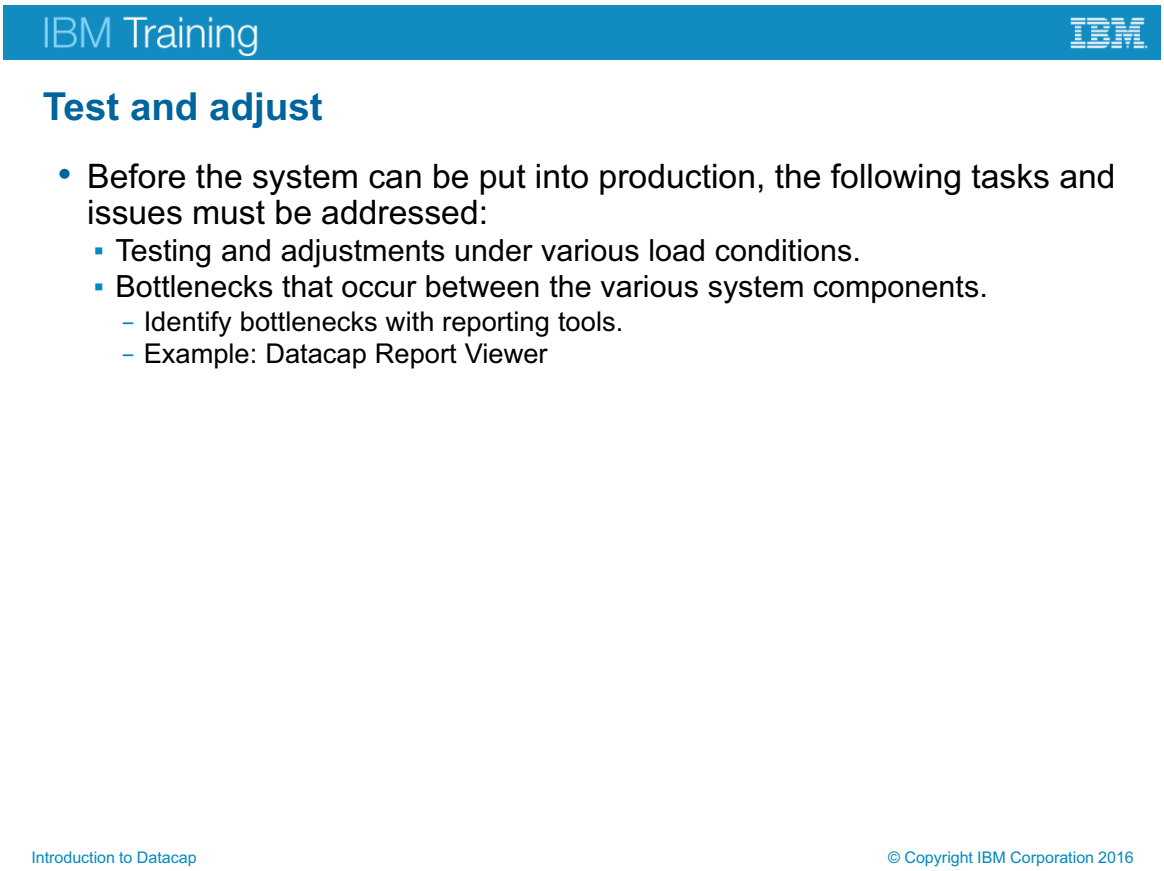


Figure 1-113. Test and adjust

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.

IBM Training

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

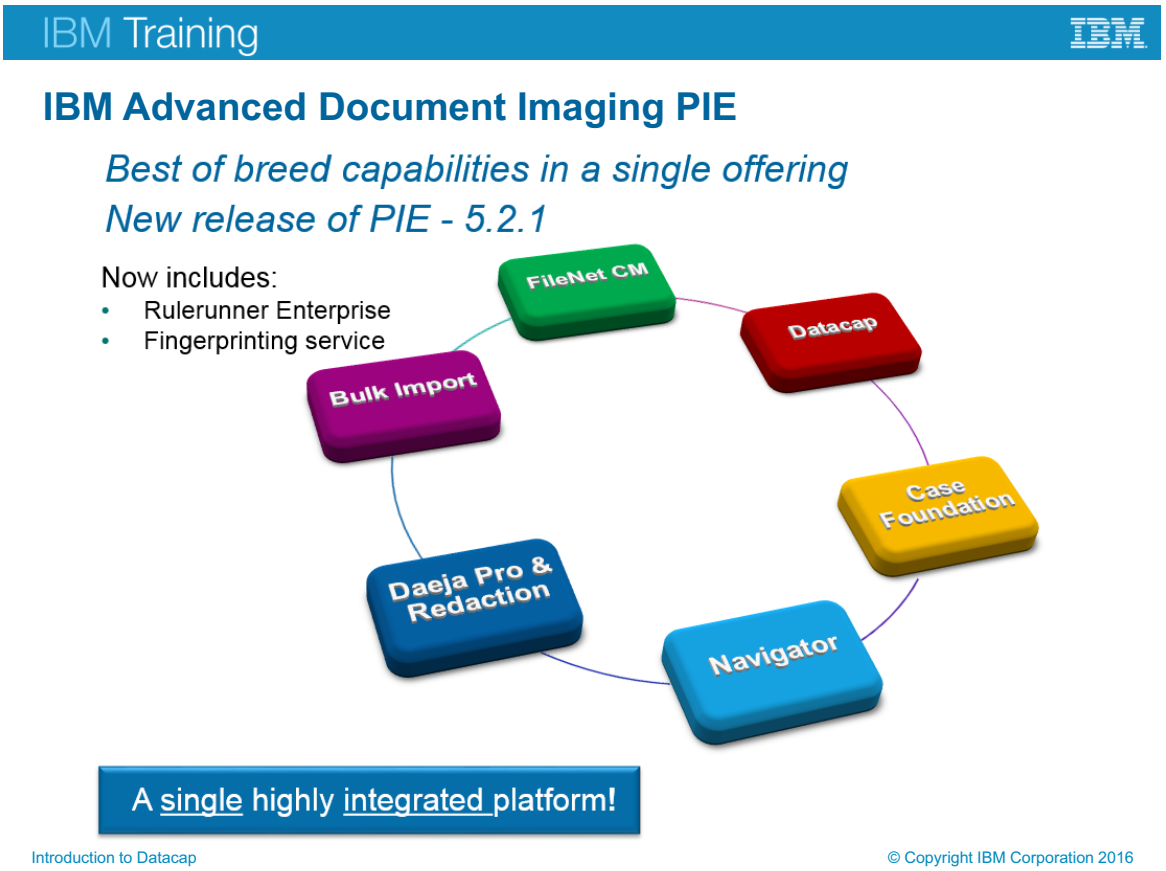


Figure 1-115. IBM Advanced Document Imaging PIE

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.

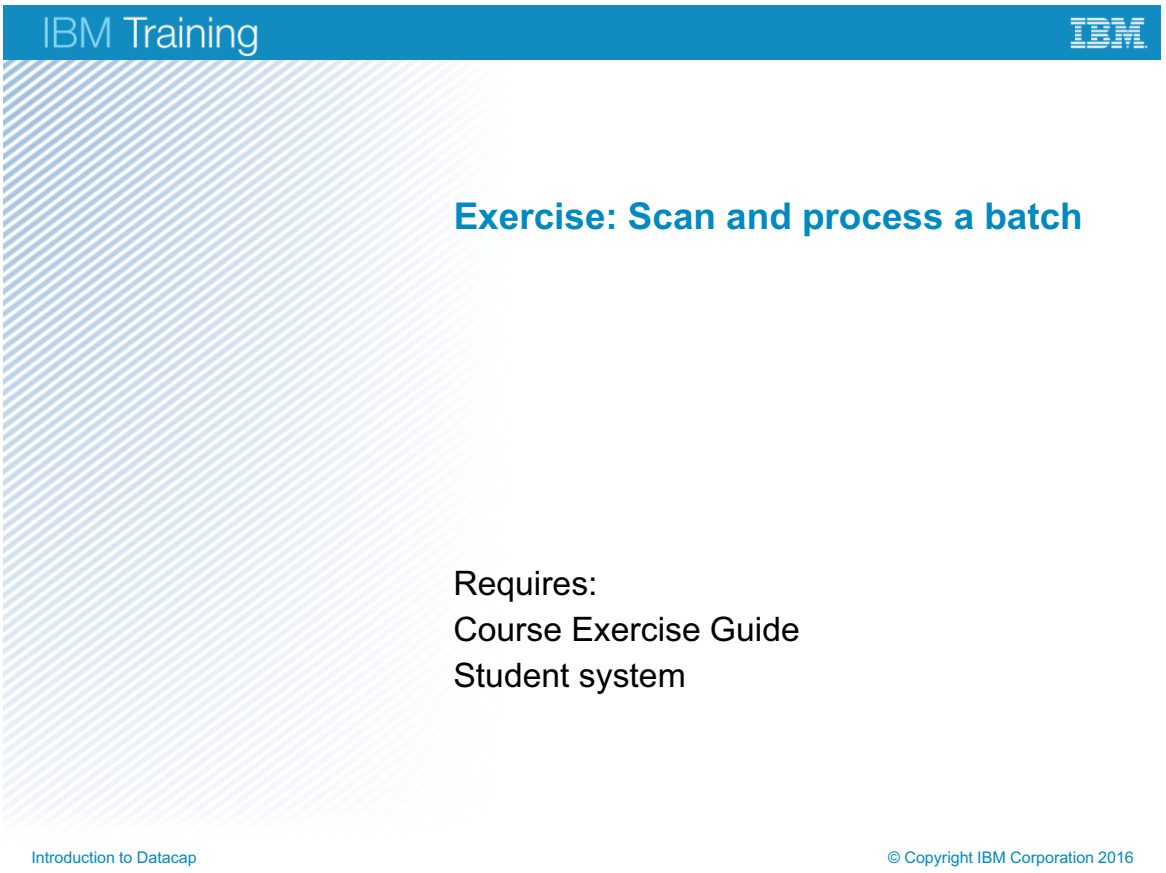


Figure 1-116. Exercise: Scan and process a batch

IBM Training

Exercise objectives

- Scan and process a batch.

Figure 1-117. Exercise objectives

Lesson 1.8. Introduction to Datacap Navigator

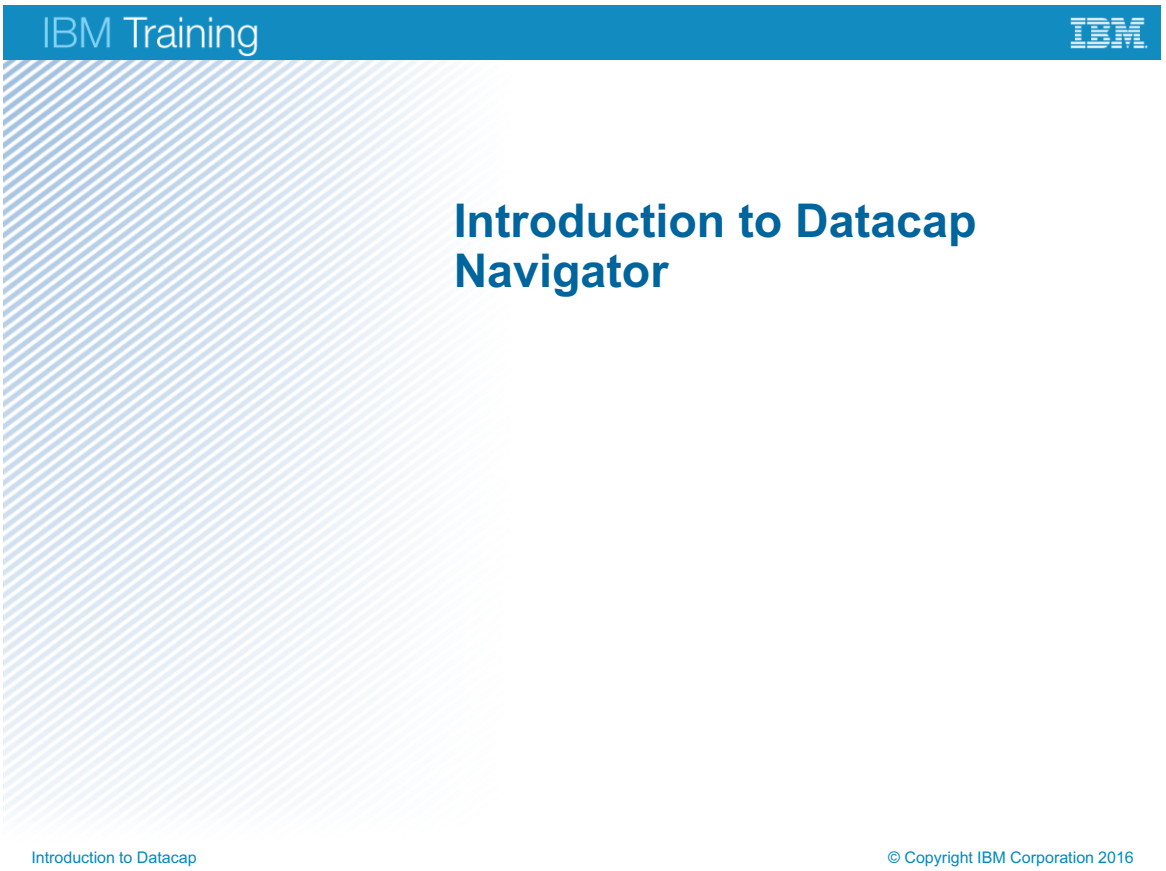


Figure 1-118. Introduction to Datacap Navigator

Figure 1-119. Lessons

IBM Training

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?

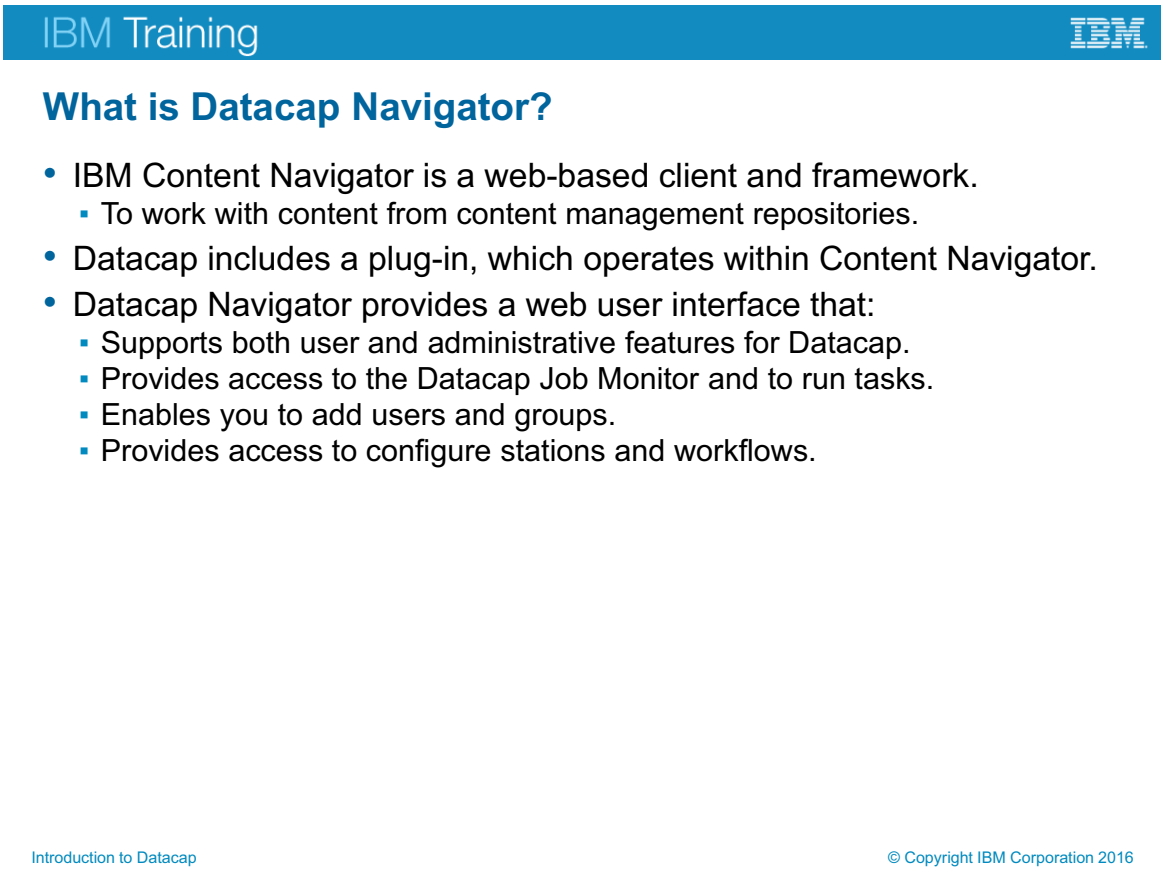


Figure 1-121. What is Datacap Navigator?

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

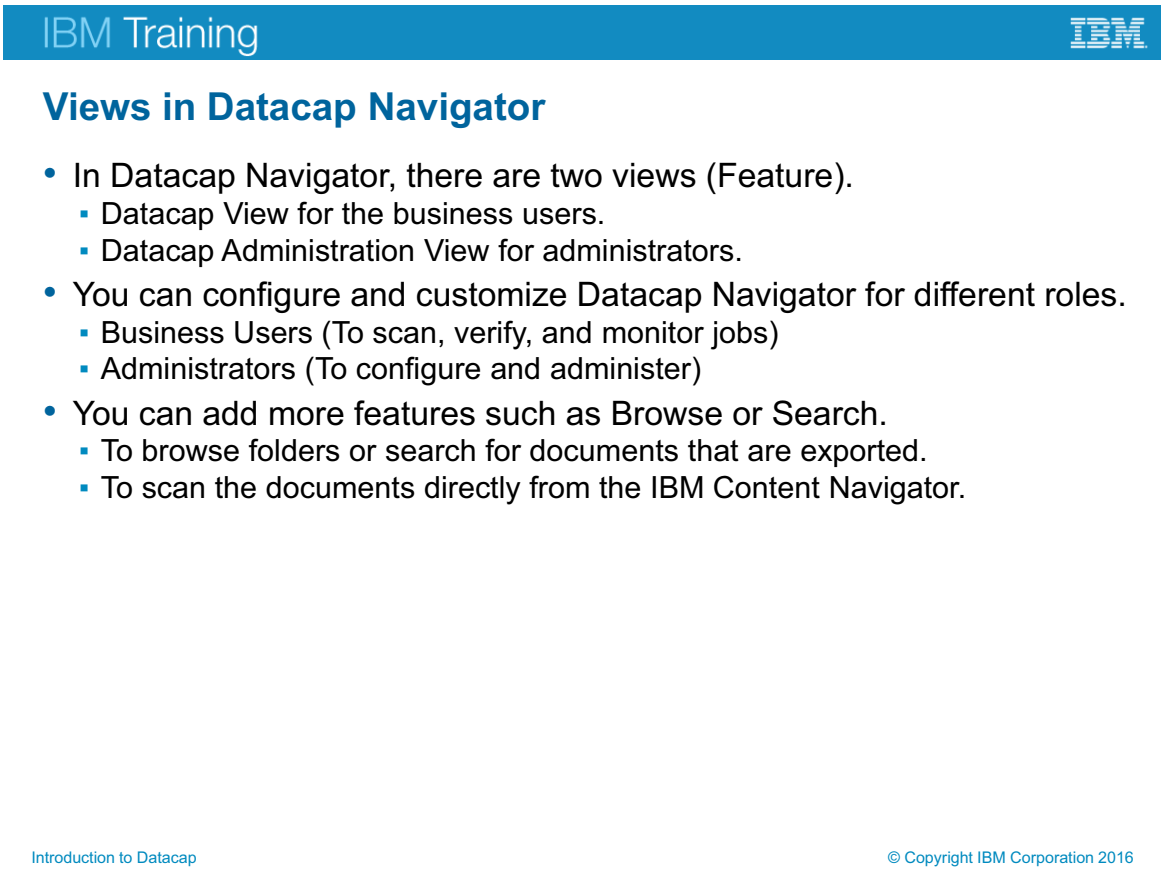


Figure 1-122. Views in Datacap Navigator

- 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.

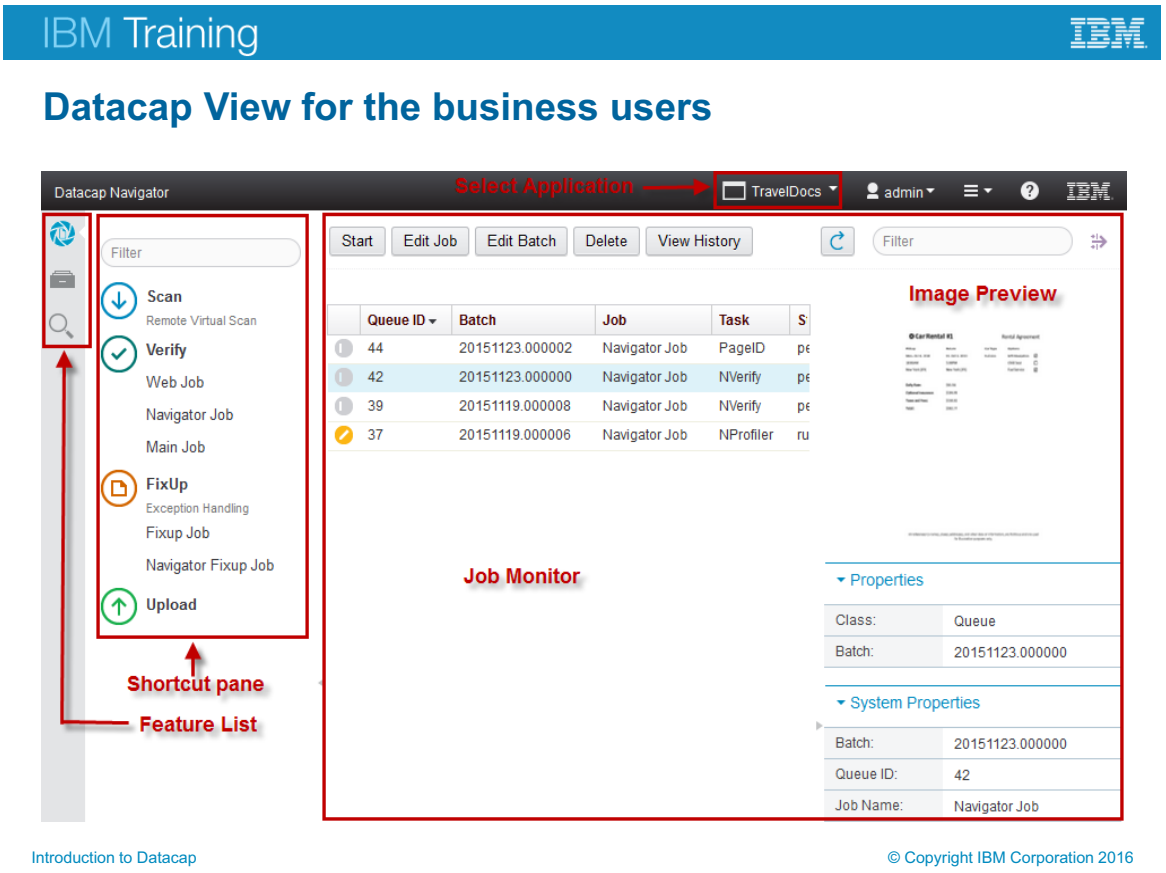


Figure 1-123. Datacap View for the business users

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

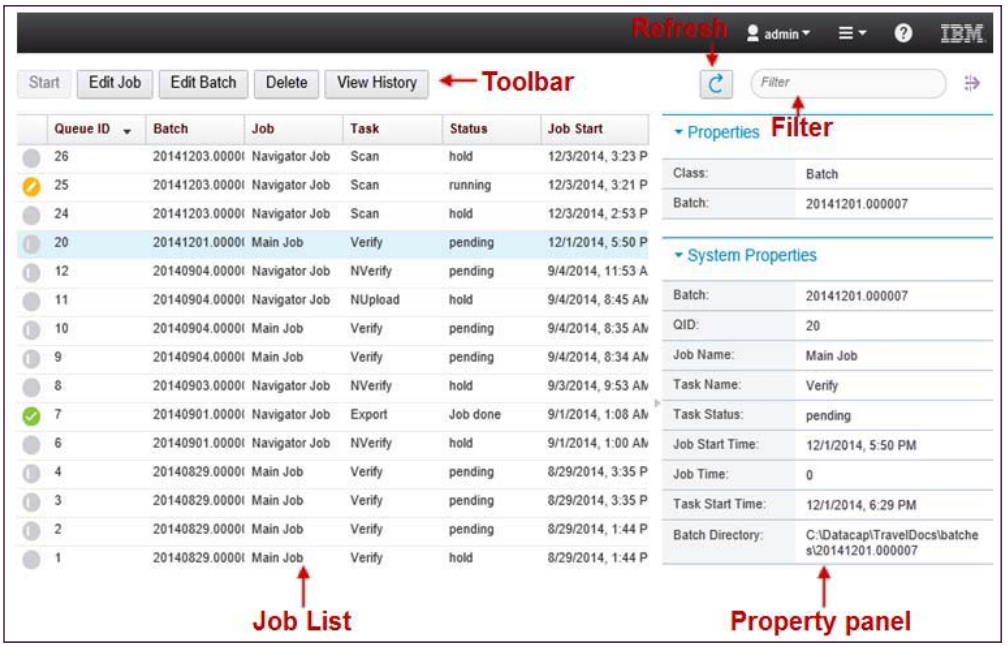


Figure 1-124. Job Monitor

- 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

IBM Training

Scan Batches

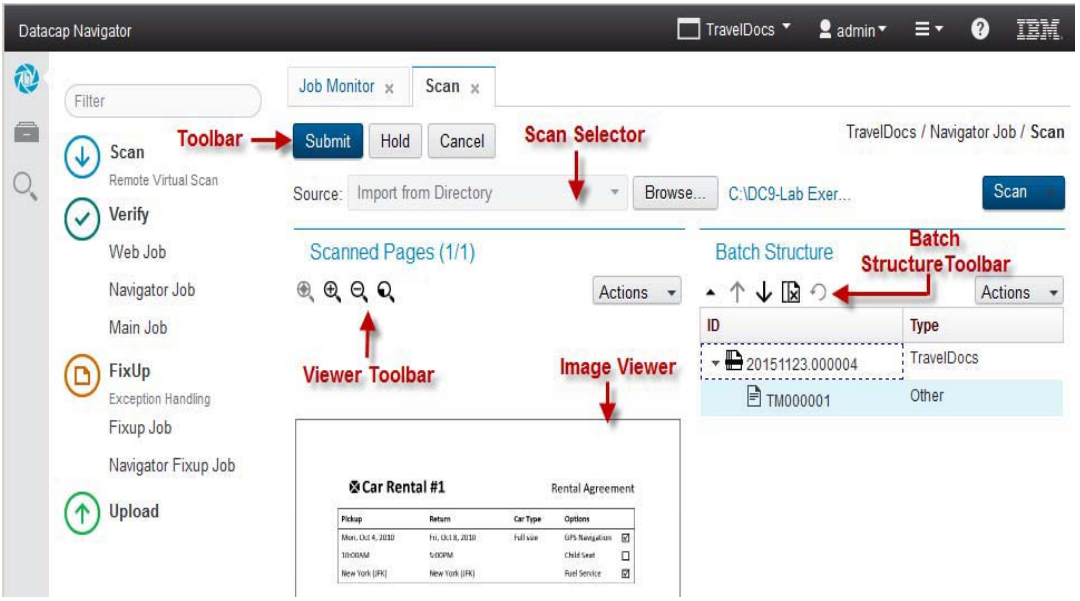


Figure 1-125. Scan Batches

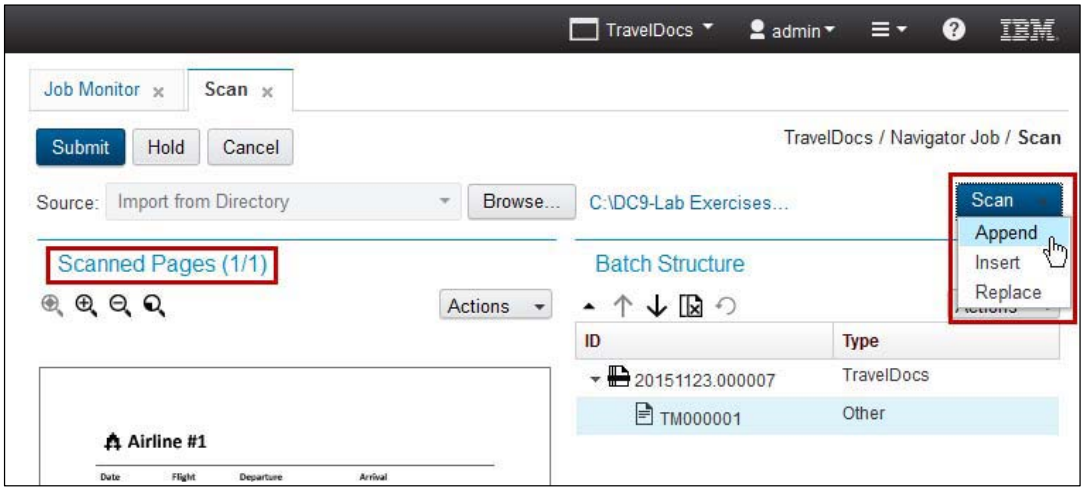
Help path

- 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

IBM Training

Rescan options in the Scan tab



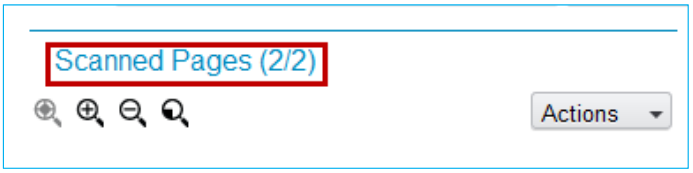


Figure 1-126. Rescan options in the Scan tab

The first screen capture shows the options for rescan.

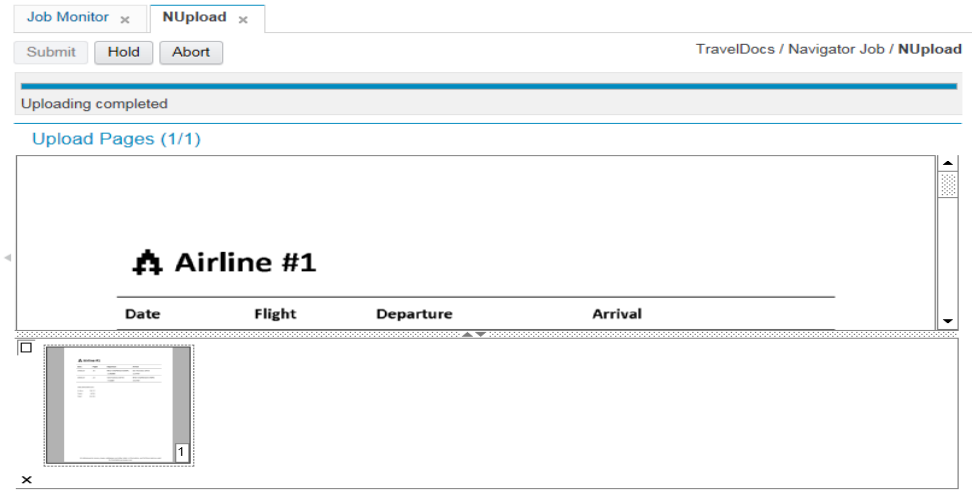
After you scan the first image, you can do a rescan: “Append”, “Insert”, or “Replace”

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

IBM Training

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.

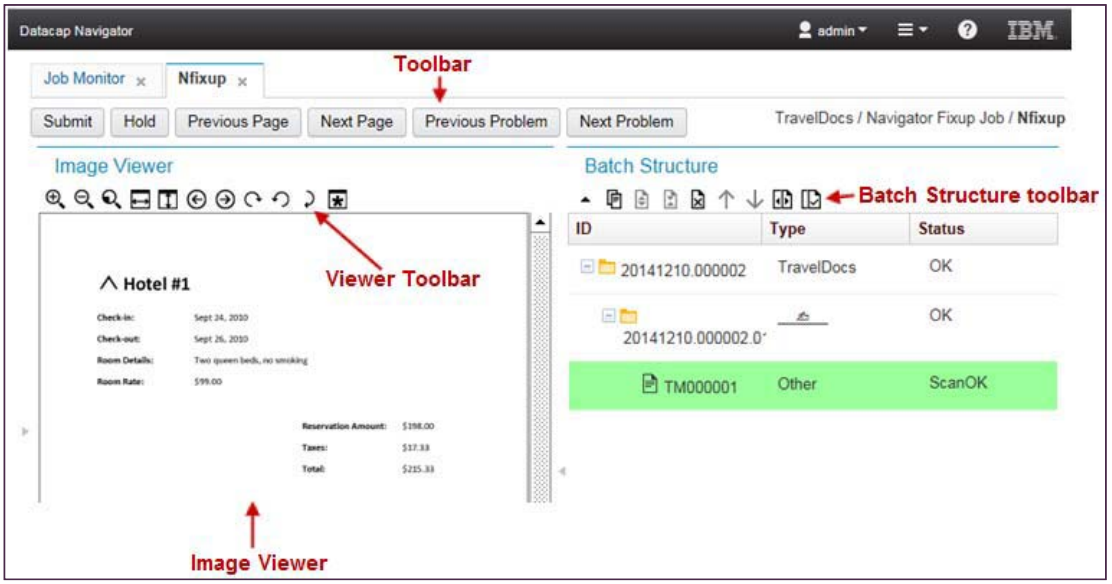
Introduction to Datacap

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

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.

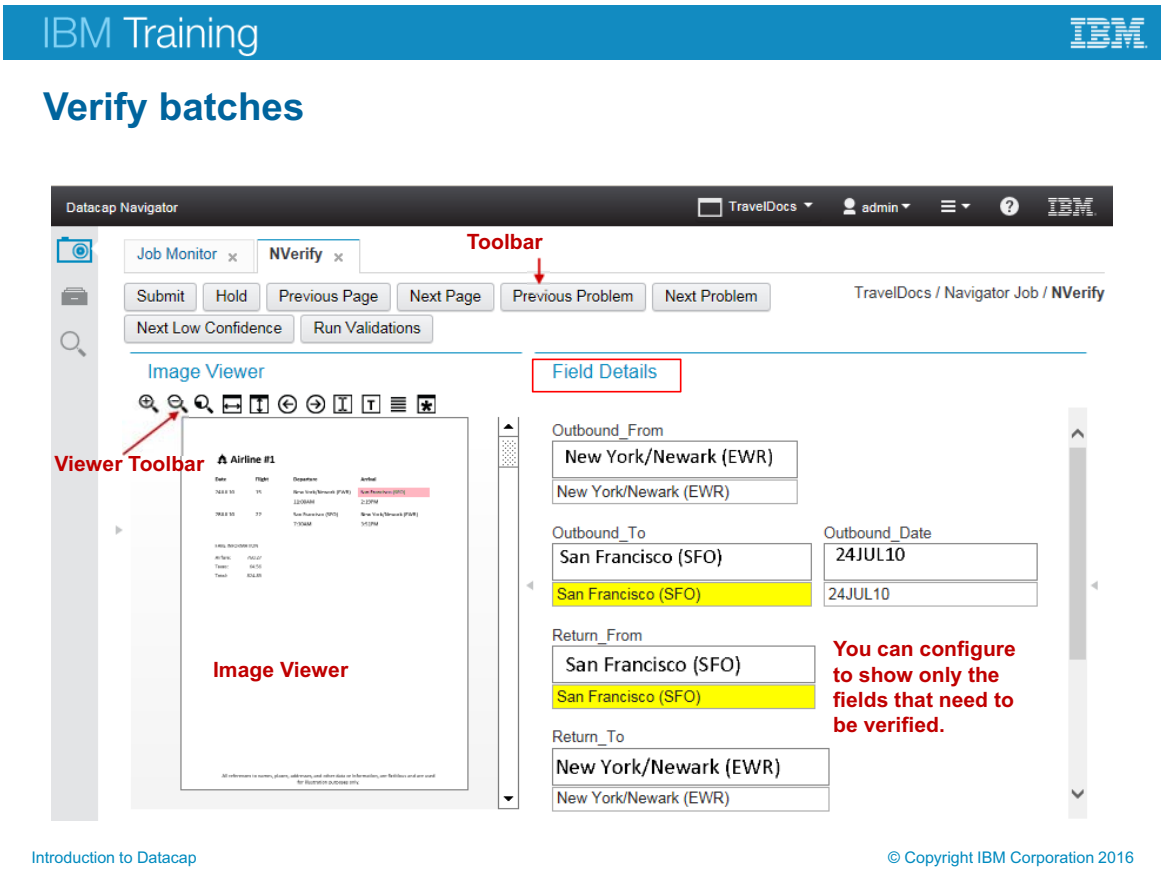


Figure 1-129. Verify batches

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

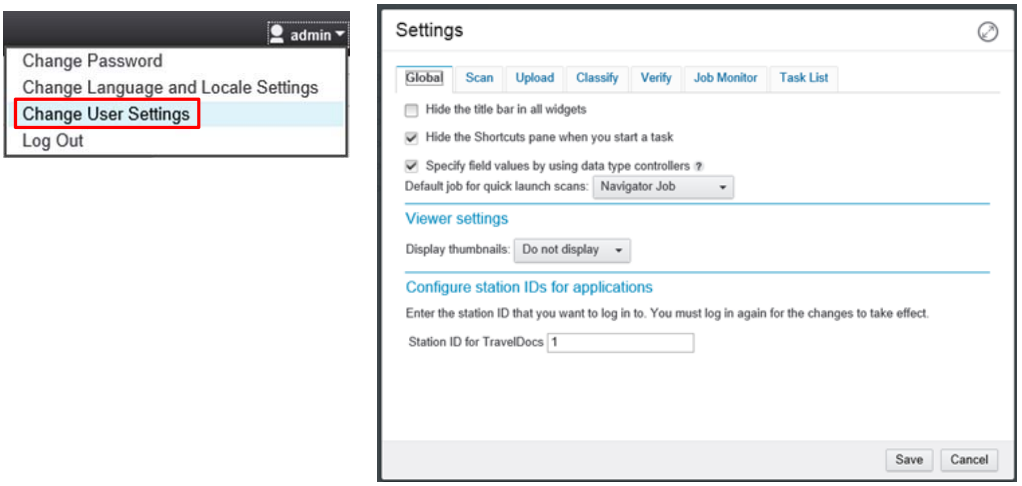
- Light pink background: fields that fail validation
- Browser side data type validation
- Support to look up data from database
- OMR – check box and radio buttons

- 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.

IBM Training

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

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

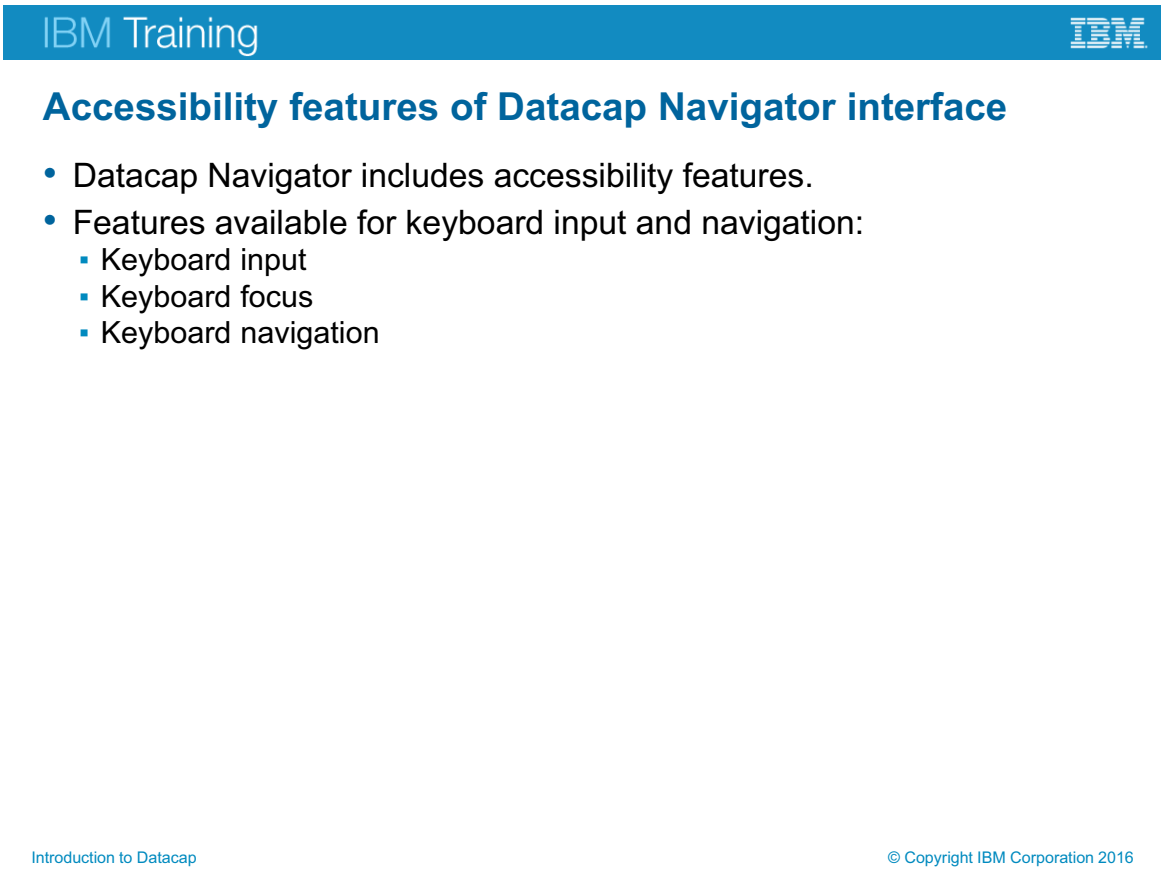


Figure 1-131. Accessibility features of Datacap Navigator interface

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/dcref003.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.

IBM Training

Demonstrations

- Explore the Datacap Navigator interface
- Scan and process a batch in Datacap Navigator

Figure 1-132. Demonstration

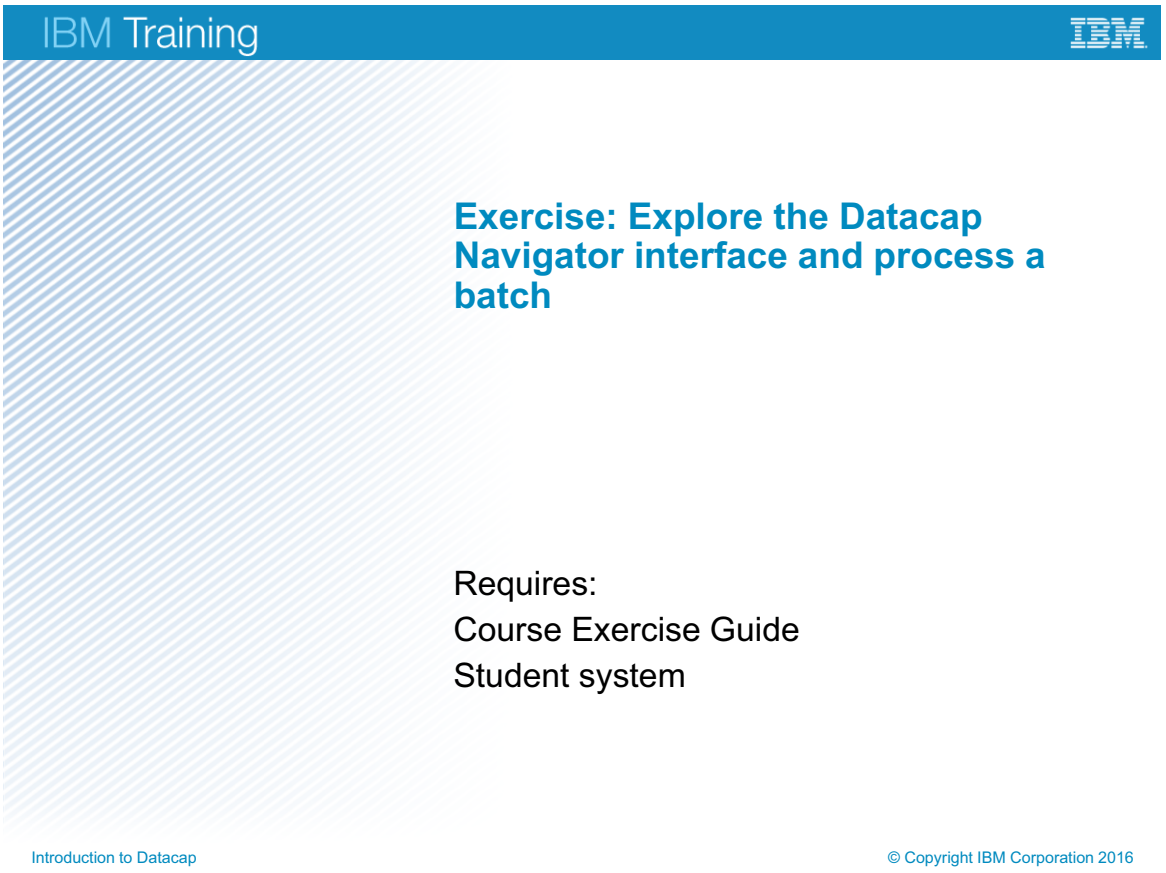


Figure 1-133. Exercise: Explore the Datacap Navigator interface and process a batch

IBM Training

Exercise objectives

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

Figure 1-134. Exercise objectives

Lesson 1.9. Datacap web client (tmweb)

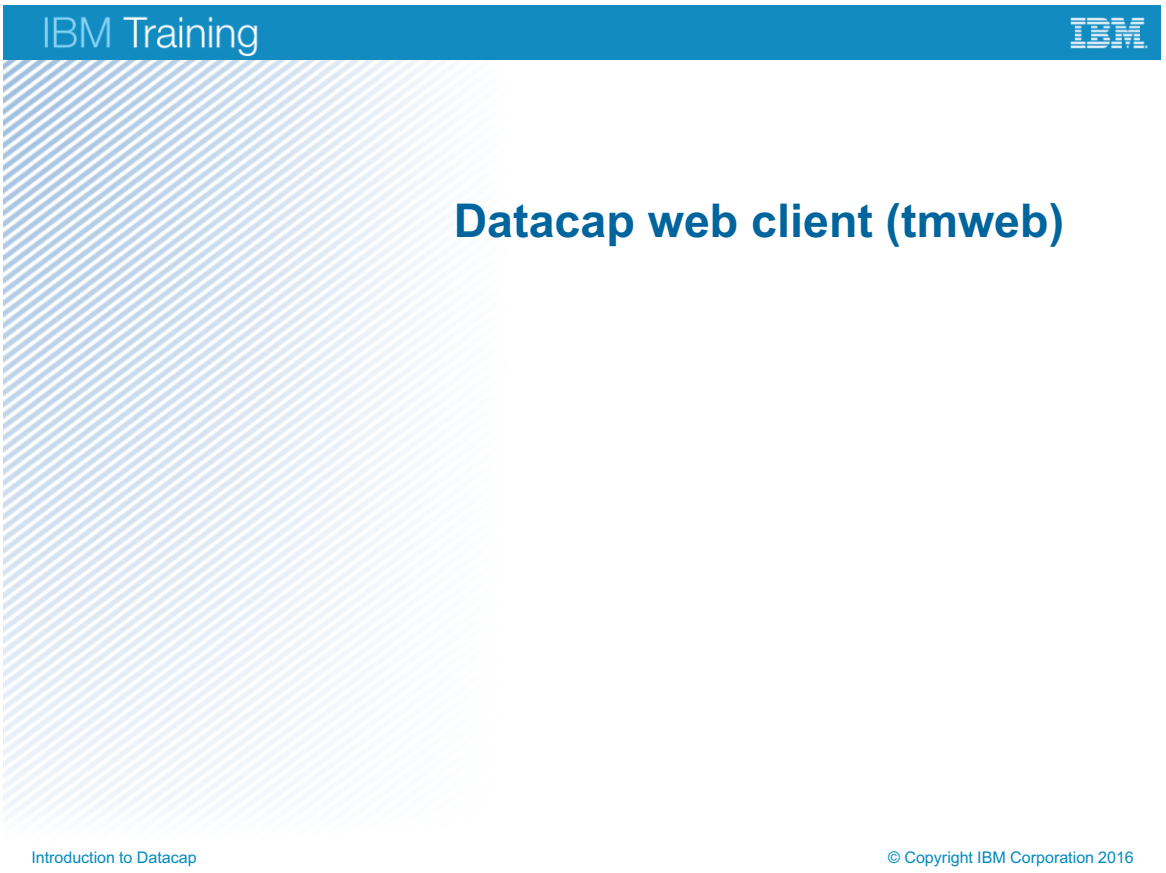


Figure 1-135. Datacap web client (tmweb)

Figure 1-136. Lessons

IBM Training

Why is this lesson important to you?

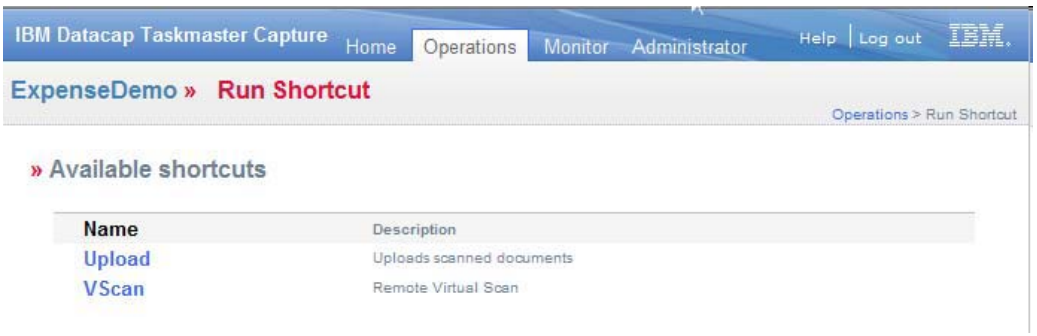
- As an administrator of an IBM Datacap capture system, you must be familiar with all configuration tasks for a functional IBM Datacap 9.0 system.
- As an application builder you must know how to use the capabilities of the Datacap Web client. Datacap Web is used for configure workflow jobs and tasks, 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 1-137. Why is this lesson important to you?

IBM Training

Datacap Web Client

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



Introduction to Datacap

Figure 1-138. 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

IBM Training

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 scancl.aspx)

Figure 1-139. 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.

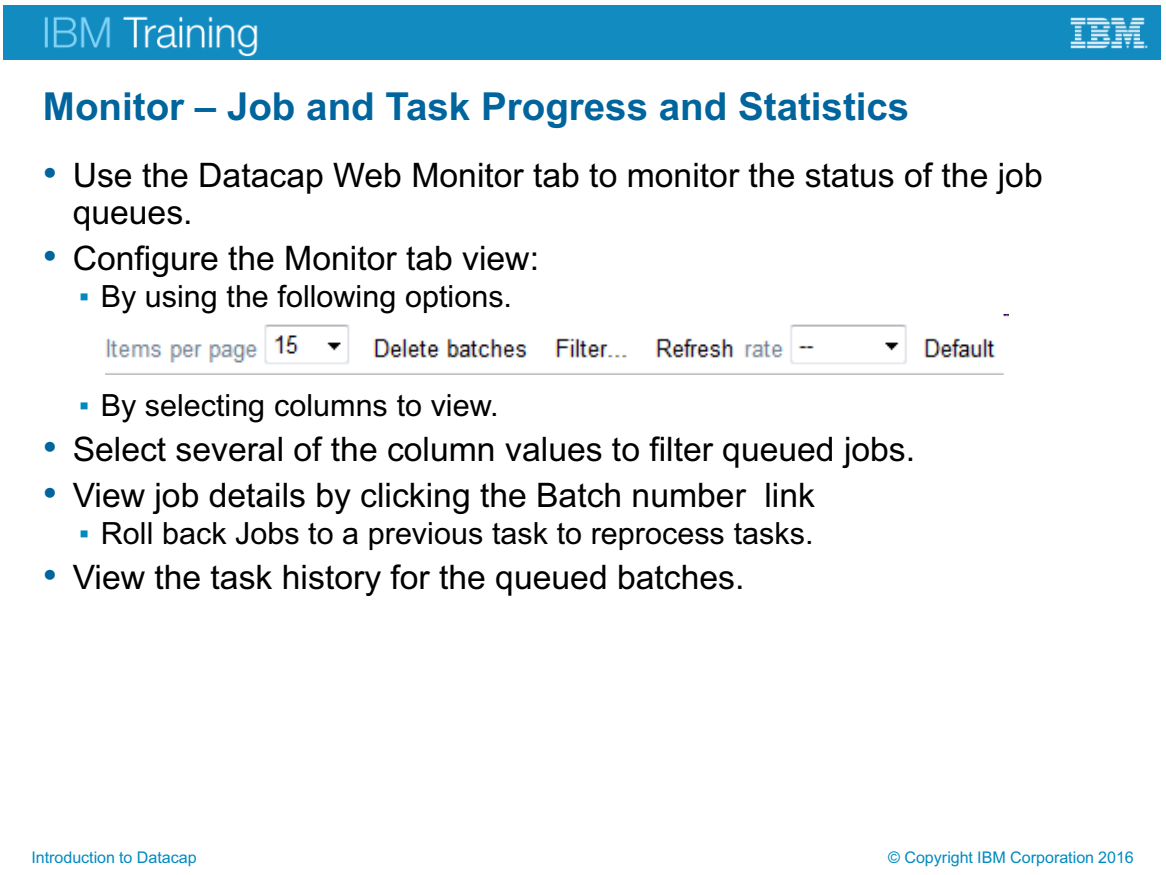


Figure 1-140. Monitor – Job and Task Progress and Statistics

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.

IBM Training

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 1-141. Administrator – Configuration

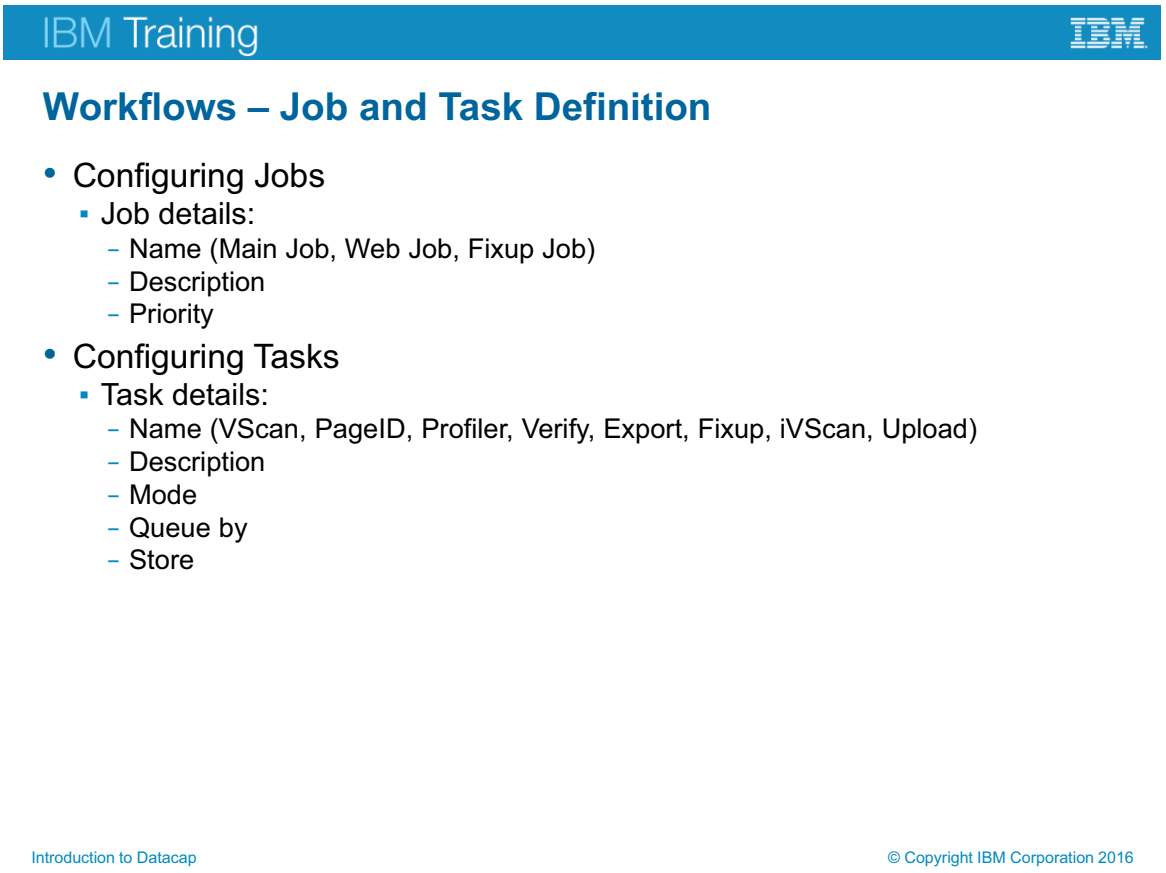


Figure 1-142. Workflows – Job and Task Definition

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.

IBM Training

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 1-143. Workflows – Task Setup

Program Value Menu Options

- **Application Clients**
  - Datacap Desktop
  - FastDoc
  - Mobile Capture
- **Background**
  - Rulerunner
- **Datacap Web Pages**
  - Scancl.aspx – Hardcopy remote scan with ISISScan or TWAINScan
  - VScancl.aspx – Virtual Scan of graphic images.
  - UplBFcl.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.
  - ProtoID.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

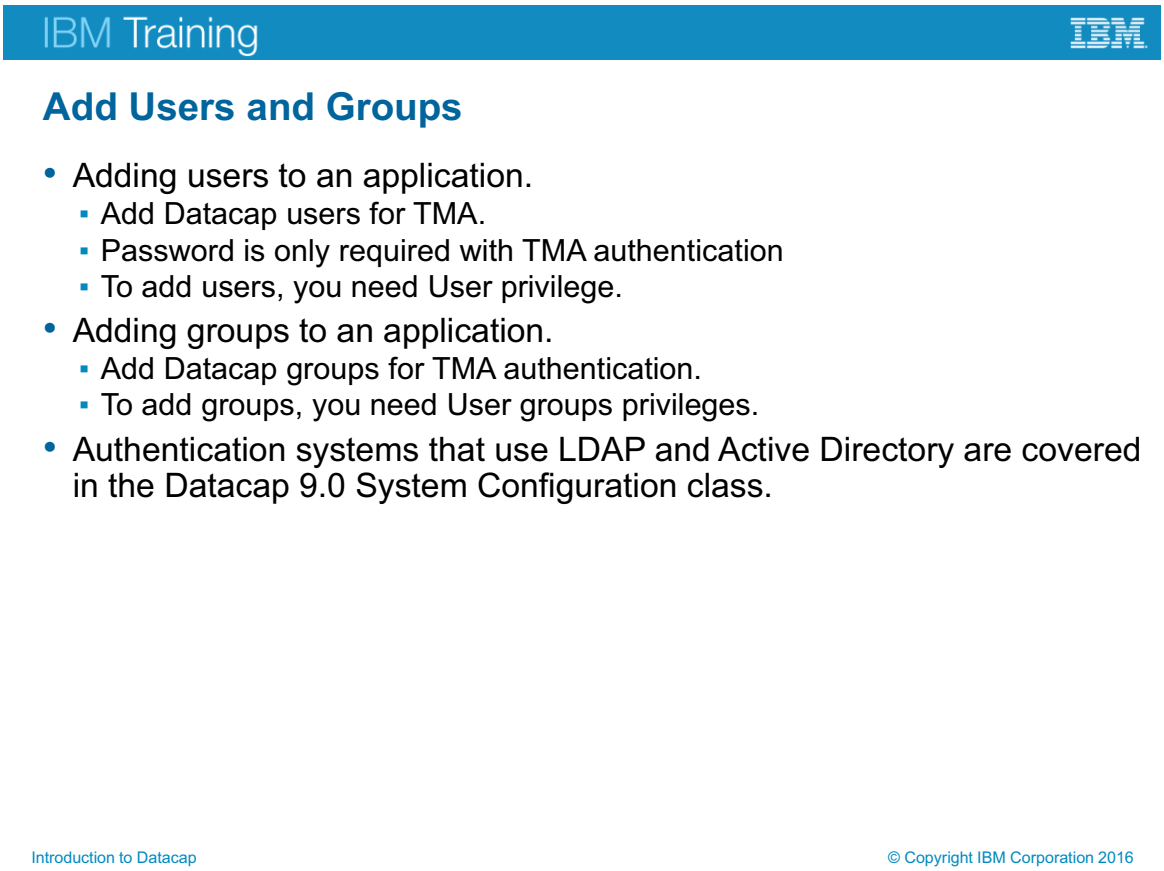


Figure 1-144. Add Users and Groups

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.

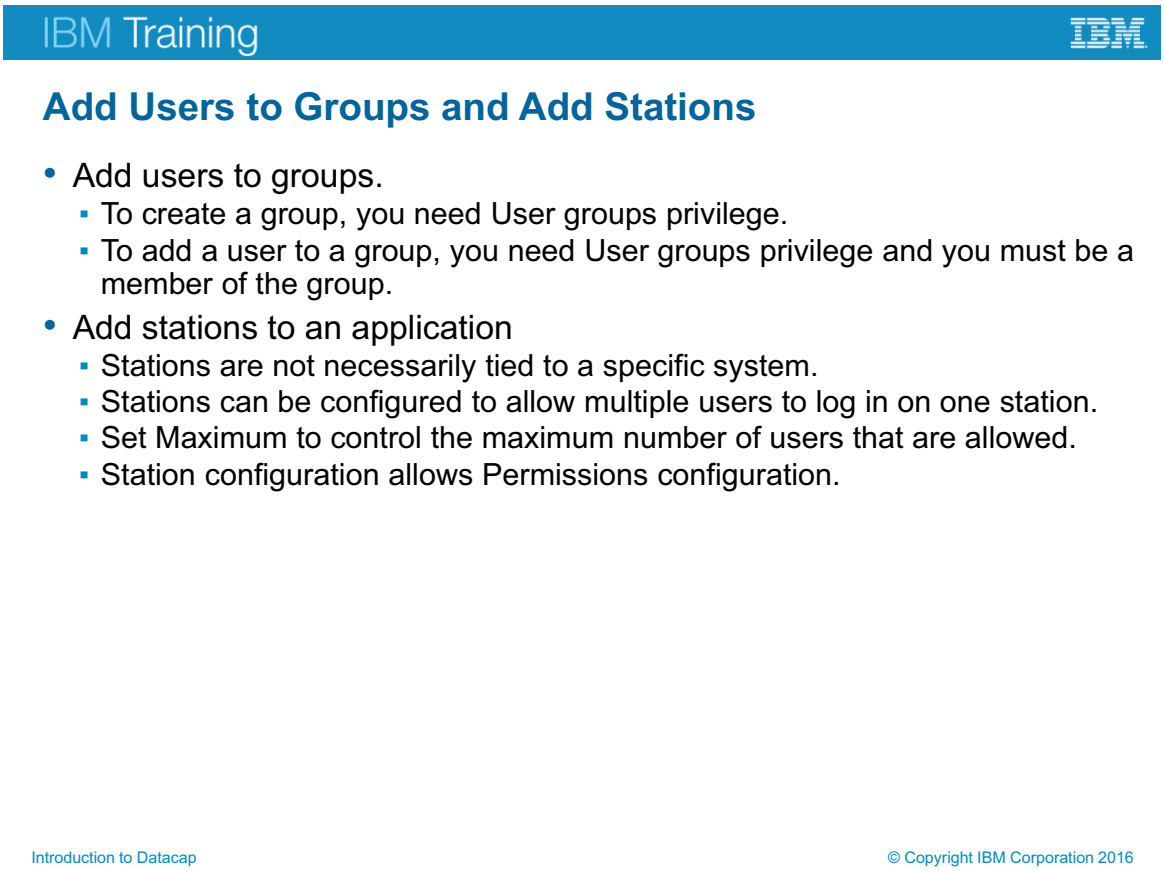


Figure 1-145. Add Users to Groups and Add Stations

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 in 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.

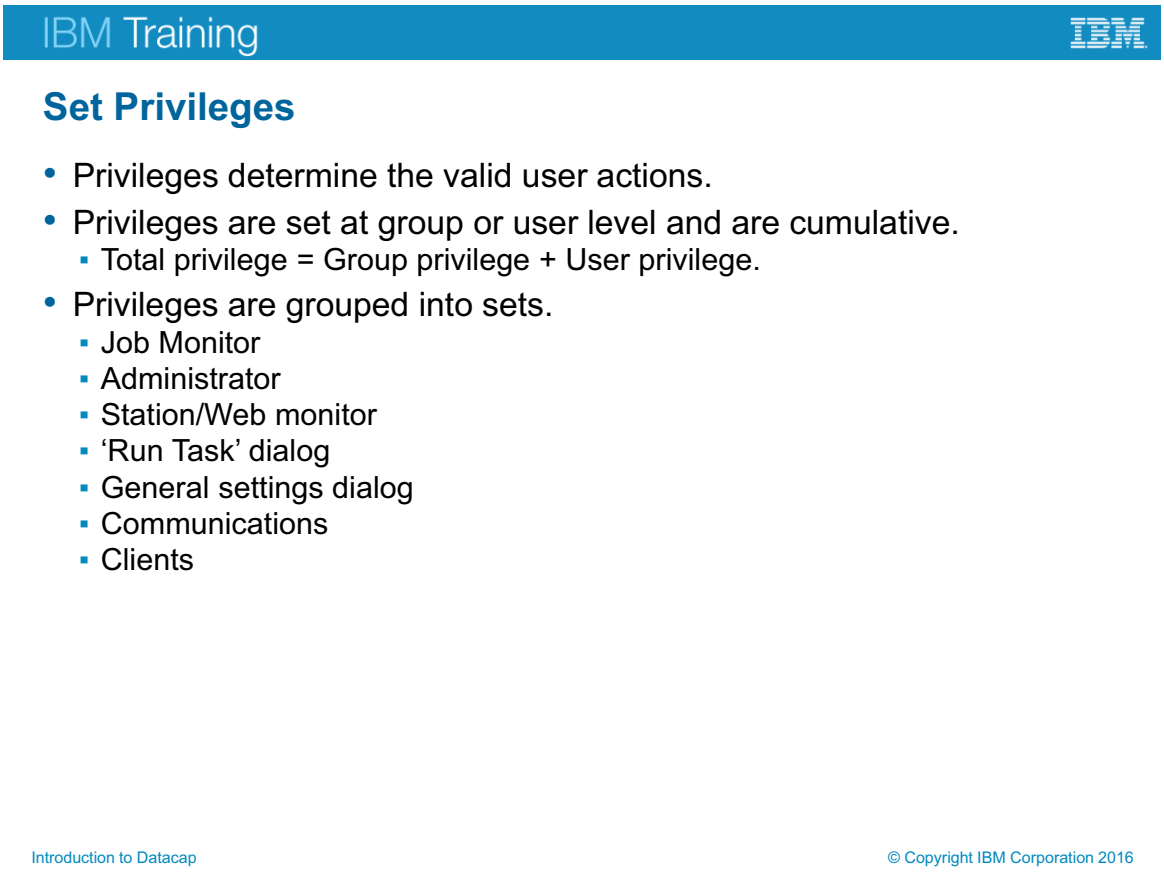


Figure 1-146. Set Privileges

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.

IBM Training

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 1-147. Set Permissions

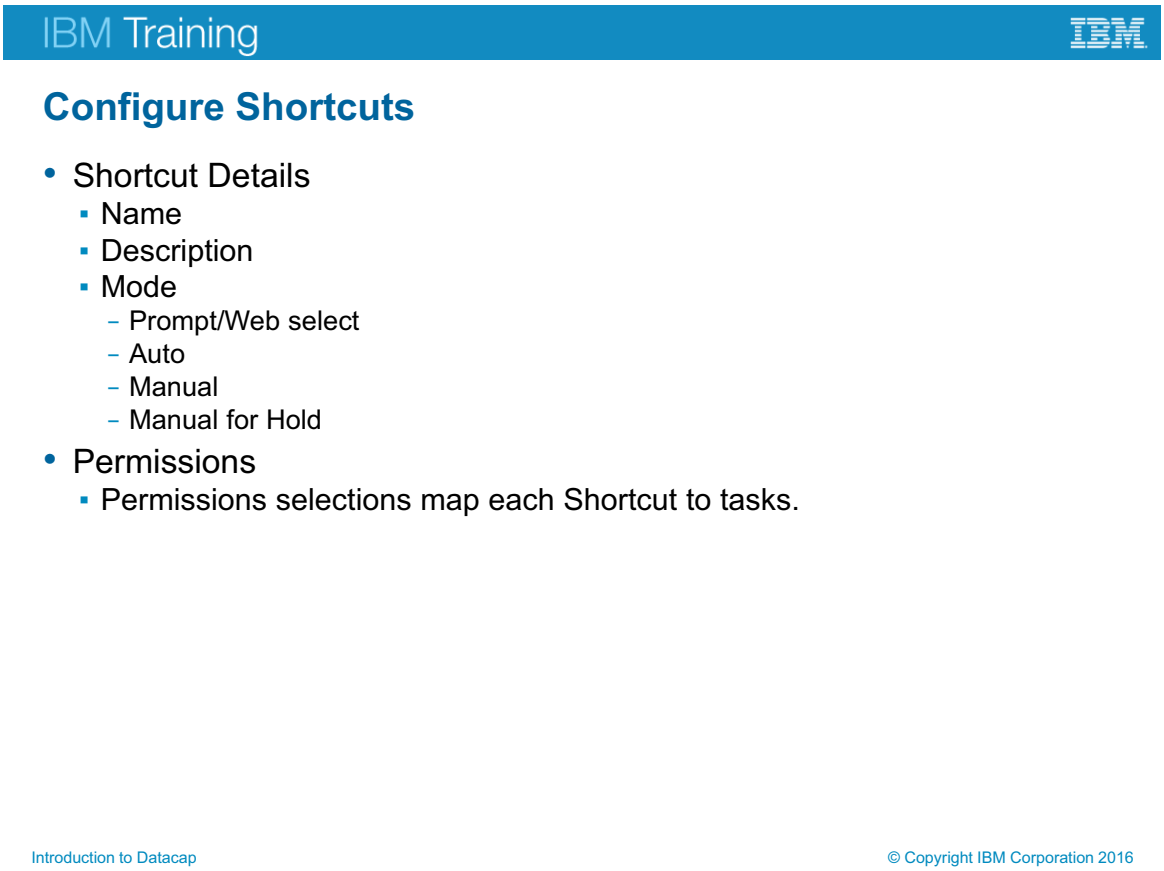


Figure 1-148. Configure Shortcuts

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.

IBM Training

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 1-149. Review questions (1)

IBM Training

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 1-150. Review answers (1)

IBM Training

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 1-151. Review questions (2)

IBM Training

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 1-152. Review answers (2)

IBM Training

Demonstrations

- Datacap Web Client Interface - Operations
- Datacap Web Client Interface - Monitor
- Datacap Web Client Interface - Admin

Figure 1-153. Demonstration

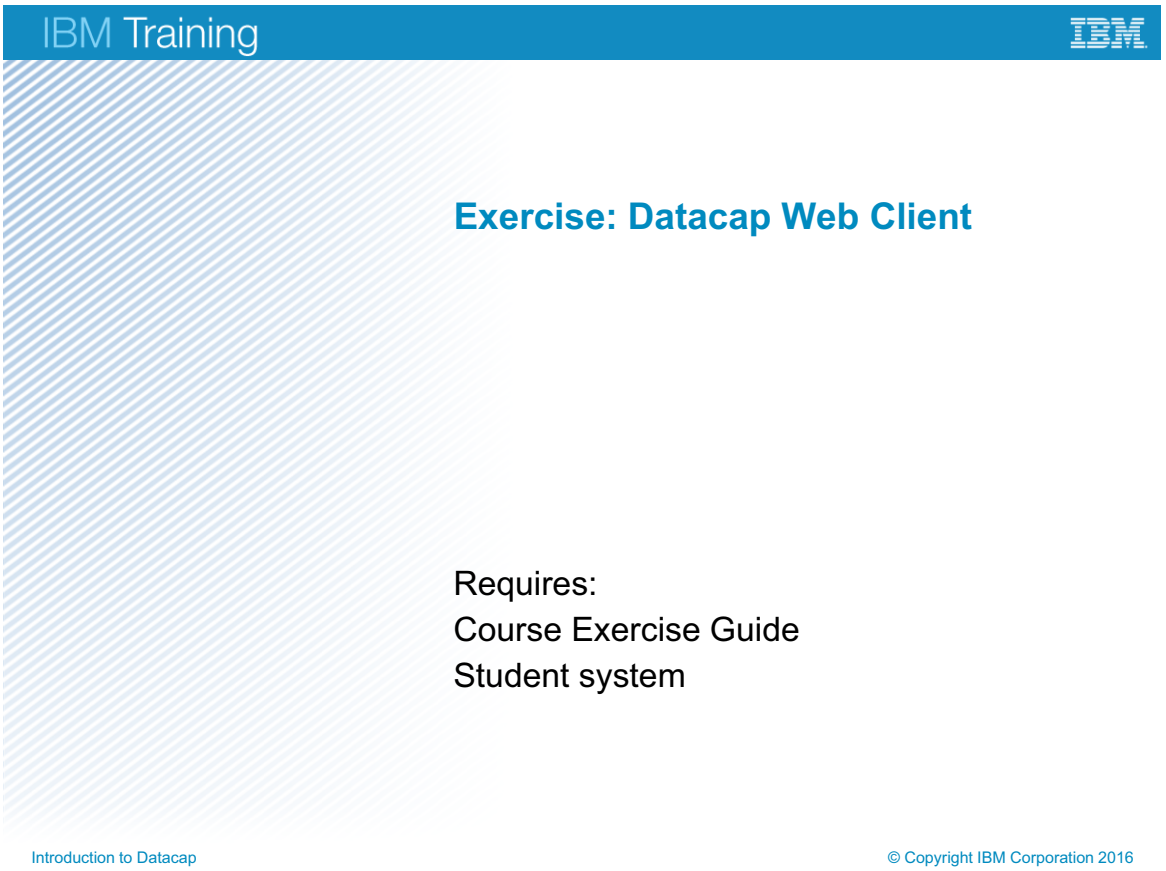


Figure 1-154. Exercise: Datacap Web Client

IBM Training

Exercise objectives

- Explore the Datacap Web Client interface

Figure 1-155. Exercise objectives

IBM Training

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 1-156. Unit summary

IBM Training