2

Overview of Datacap Studio

This overview of Datacap Studio covers the following topics:

2.1 Datacap Studio Architecture

Datacap Studio plays a central role in setting up and testing an application that utilizes Rulerunner Service.

✓ The basic diagram on the next page highlights the essential components that you manage with Datacap Studio.

At the top:

- **Datacap Studio** prepares and manages a **Document Hierarchy** with four processing levels – **Batch**, **Document, Page** and **Field.** The Document Hierarchy clearly identifies and distinguishes **objects** at each level. Each object has **properties** (Chapter 5 explains this core component.) At the top of the Document Hierarchy, the static **Application** object identifies the application.

- Datacap Studio is responsible for defining an application's **fingerprint(s).** Each fingerprint is a pair of files that represent a particular **type** of **Page**, or in some cases a particular version or variant of a page type – a page with data that is to be recognized, verified, validated and exported.
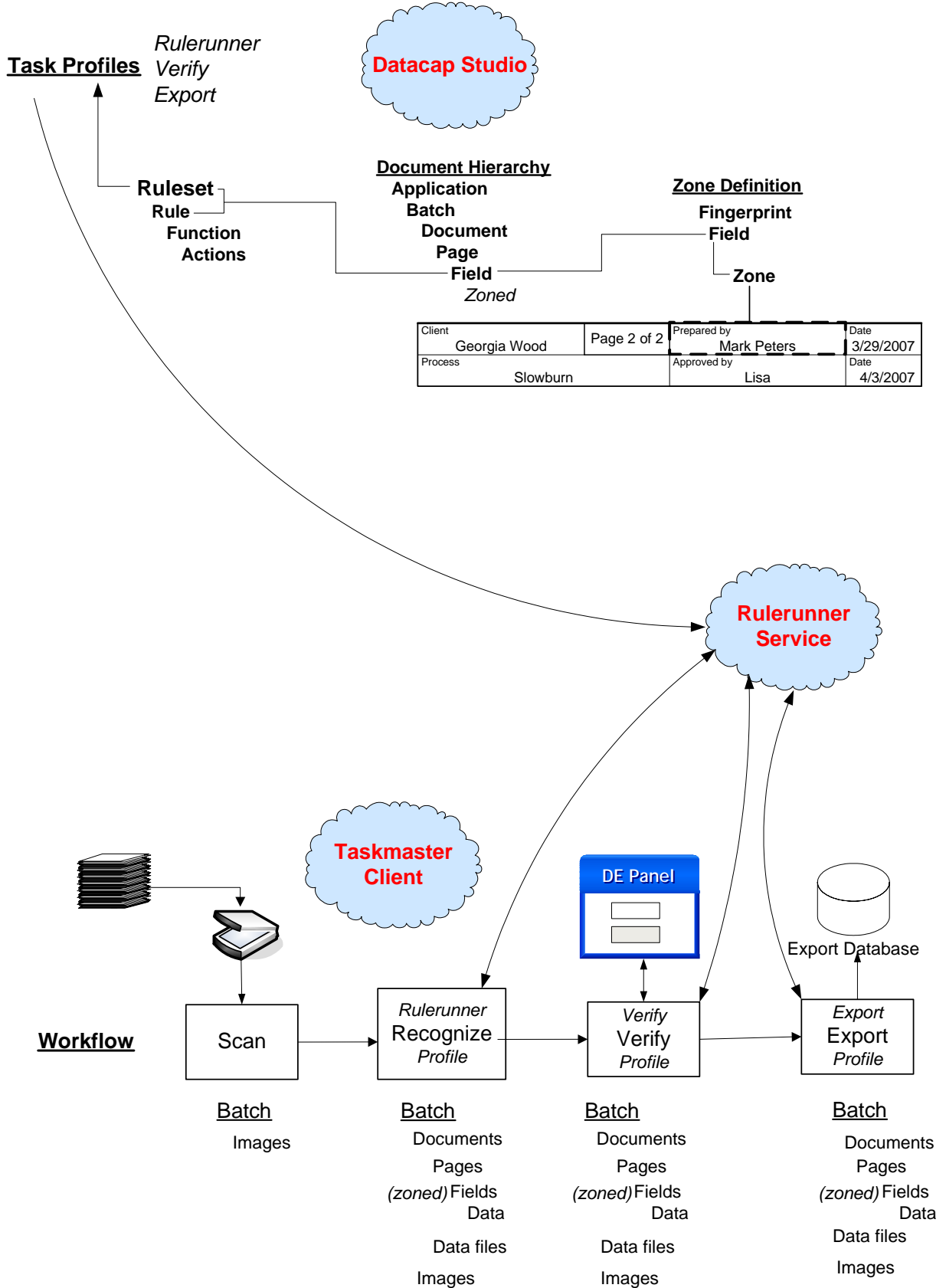
    One member of the pair is an Image file (.tif);

    The other is a Recognition file (.cco).

    For more about fingerprints, see Chapter 7.

- **Zoning** locates fields and blocks of fields on the Image file. This procedure can also assign special recognition criteria to individual fields.
  The diagram highlights one zoned **Field**: The Document Hierarchy includes this **Field** object, its zone(s) and additional properties.

- A **Rule** consists of one or more **Functions**, and a function is made up of one or more **Actions**. In addition, a rule belongs to a particular **Ruleset.** The resulting **Ruleset/Rule** combination is **bound** to a specific object of the Document Hierarchy – in this case, to the **Field** object of the Document Hierarchy.

- An application's *Rulerunner* tasks (Recognize, Verify and Export tasks, for example) run or "execute" the rules of one or more appropriate RuleSets defined by their **Task Profiles** (Chapter 6). Datacap Studio assigns specific Rulesets to Task Profiles**.** During the set up of a workflow's **rulerunnner tasks**, a Task Profile is associated with each rulerunner task.

At the bottom of the chart, the application's **Taskmaster Client** carries out the tasks of a workflow. In this basic example, the initial Scan task creates a new batch, captures images of paper documents using an image scanner, and adds the resulting Image files to the batch. Because the Scan task is *not* a rulerunner task, it is not linked to one of Datacap Studio's Task Profiles.

Task Profiles

Rulerunner
Verify
Export

Datacap Studio

Ruleset
  Rule
    Function
      Actions

Document Hierarchy
  Application
    Batch
      Document
        Page
          Field
          Zoned

Zone Definition
  Fingerprint
    Field
      Zone

Client		Page 2 of 2	Prepared by		Date
Georgia Wood			Mark Peters		3/29/2007
Process			Approved by		Date
Slowburn			Lisa		4/3/2007

**Rulerunner
Service**

**Taskmaster
Client**

DE Panel

Export Database

Workflow

Scan

Rulerunner
Recognize
Profile

Verify
Verify
Profile

Export
Export
Profile

<u>Batch</u>

Images

<u>Batch</u>

Documents

Pages

(zoned) Fields

Data

Data files

Images

<u>Batch</u>

Documents

Pages

(zoned) Fields

Data

Data files

Images

<u>Batch</u>

Documents

Pages

(zoned) Fields

Data

Data files

Images

✓ In this example, the workflow's other tasks – Recognize, Verify, and Export – *are* Rulerunner tasks: each instructs Rulerunner Service to perform the Rulesets assigned to the task's Task Profile. Setting up a rulerunner task involves associating the appropriate Task Profile to a Task Definition. (For details, see Chapter 6.)

Rulerunner Service is the entity in the middle. Rulerunner Service receives information about the task and the batch from the rulerunner task, runs the Task Profile (RuleSets, Rules, Functions, and Actions) on the batch, and returns the updated batch and related information to the Rulerunner task.

2.2 Installation and Setup of Datacap Studio

The Datacap Taskmaster Client Installation automatically installs Datacap Studio. For complete installation and configuration instructions, see the Taskmaster Installation and Configuration Guide.

2.2.1 Datacap Studio Files

**Taskmaster Client Installation** installs files used by Datacap Studio:

- Datacap Studio's configuration and executable files are in the **Datacap** directory's **DStudio** folder.

- Application files for sample and foundation applications are in individual sub-folders of the **Datacap** directory.

- Individual application sub-folders of the **Libs** directory of the **RRS** folder have Rules files (.rul) with information about an application's rules, and a single collection.xml file that point to the application's Taskmaster databases, process and batches folders, Document Hierarchy and Task Profiles.

- The **RRX** sub-folder of the **RRS** folder contains Action Library files (.rrx) that are shared by all applications.

✓ *Important!* Datacap Studio can work with an application only if its DStudio files are in the **Libs** sub-folder of the **Datacap** directory's **RRS** folder. Refer to the Taskmaster Installation and Configuration Guide for complete setup instructions.


Datacap Directory

Here is a closer look at certain folders of the **Datacap** directory. Note that the files and folders will vary depending on which Taskmaster component was installed (Client, Server, Web, etc.).

Every *Taskmaster* Client application has the folders below. When you create a new application, application folders are inserted into the **Datacap** directory automatically. Chapter 4 explains this process, as well as the nature and locations of the application's processing files.
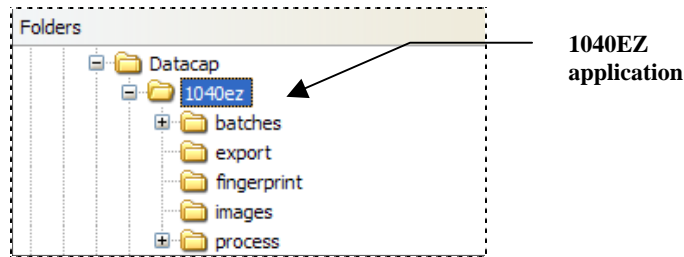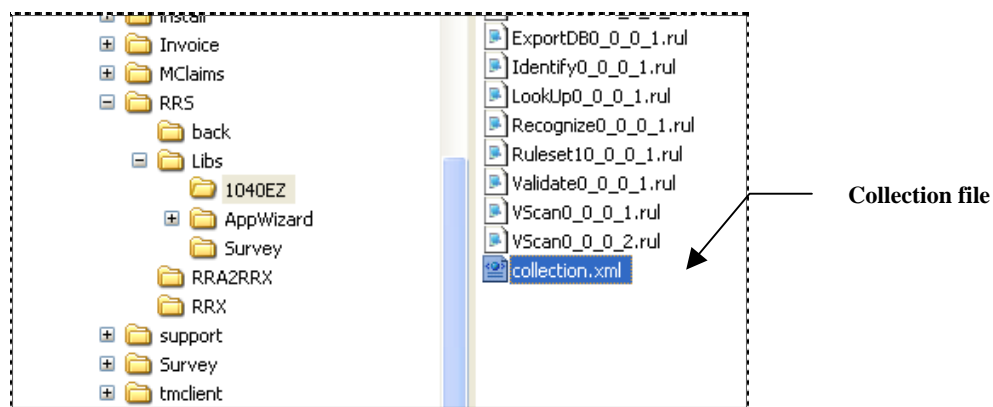

1040EZ Application

Further down, the **Datacap** directory's **RRS** folder has a **Libs** sub-directory with another set of application folders. Each folder contains Rule files (.rul) and a Collection file (.xml) with details of the Ruleset/Rule combinations that govern task procedures.
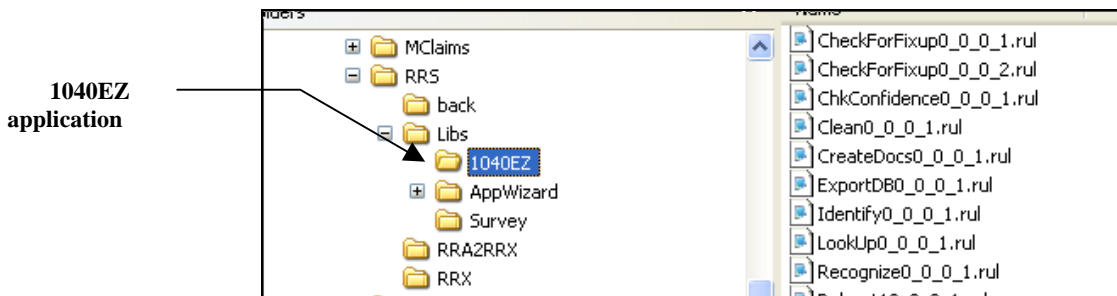
1040EZ Ruleset/Rule Combinations

1040EZ Collection File

The first illustration below shows a portion of the Ruleset/Rule information in the Collection file; the second shows you how the file lists Task Profile specifications.

Ruleset/Rule Information

Task Profile Information

Finally, the **Datacap** directory's **RRX** folder contains Libraries (.rrx) filled with the actions that power rules.