

# Rule Manager Reference

---

## Glossary

This Glossary contains brief definitions of terms associated with *Taskmaster*, a *Taskmaster* application – and *Rule Manager*.

### Action

A concise procedure that is part of a rule. Together, a rule's actions determine how a particular task is to deal with a specific object of the Document Hierarchy...a particular page or field, for example.

An action belongs to one of a number of categories such as **Locate**, **Validate** and **LookUp** – and are stored in Action files (.rra).

A rule can use actions from any category that is part of the RuleSet Type's Actions Library. *See also: Rules, RuleSet Type, Actions Library, Document Hierarchy.*

### Actions Library

The combination of Action files (.rra) that hold actions available to rules of a specific RuleSet Type. Actions that belong to a particular category – **Locate** actions, for example. A RuleSet Actions Library contains one or more files.

### Admin Database

The *Taskmaster* database that contains definitions of workflows, jobs, tasks, users and workstations. *See also: Engine Database, Rules Database, Taskmaster.*

### Administrative Privileges

Attributes of a User Definition designating the ability of an individual to access an application's administrative dialogs and data, and to carry out the administrative functions of a *Taskmaster* application. *See also: User Definition, Job/Task Permissions.*

### Application

The principal organizing entity of a *Taskmaster* configuration. An application's Document Hierarchy identifies *structural* objects at four levels: **Batch**, **Document**, **Page** and **Field**. A Workflow Hierarchy organizes the application's *processing* entities at three levels: Workflow, Job and Task. *See also: Document Hierarchy, Workflow Hierarchy*

### AutoDelete

A *Taskmaster* utility you can use to identify those batches that meet your selection criteria, then delete some or all of these batches from the application's **Batches** directory. *AutoDelete* updates tables of *Taskmaster's* Engine database with the results of the deletion and retains evidence of the deleted batch, according to guidelines you prescribe.

### Batch

A *Taskmaster* application's processing vehicle *and* the organizing object at the top of its Document Hierarchy. A workflow's opening Scan task assembles a batch when it scans one or more pages, creates an Image (.tif) file for each page, and assigns the Image files to the batch. The Scan task then adds a Batch Contents file (.xml) to the batch with details of its contents and, optionally, a Log (.log) file covering the task's performance. Each succeeding task in the workflow processes the batch formed by the Scan task—and adds files associated with that task to the batch. *See also: Workflow, Scan Task, Image File, Page, Batch Contents File, Data File, Log File.*

### Batch Number

A unique code assigned by the KScan task to identify a batch. *Taskmaster* supports various Batch Numbering formats. Also known as the Batch ID, the standard *Taskmaster* format is **YYYYMMDD.nnn**.

### Batch Pilot

The Task Development workshop for most *Taskmaster* tasks. *Batch Pilot* technology is responsible for the design, testing and implantation of these tasks – and for the maintenance of *Rule Manager* and the tools of the **Rule Manager Window**.

### Binding

The procedure that links a rule's definition to an object of the Document Hierarchy. A rule that is bound to a **Page** object, for example, determines how the task that applies the rule processes the page.

### Captured Value

The value of a target word on a current page after it has been recognized, linked to a **Field** object of the Document Hierarchy, and added to the Data file of the page. *See also: Field Object, Target Value.*

### .cco File

A Processing file that is one of two files that comprise a fingerprint: the other is the fingerprint's Image file (.tif). The Processing file (.cco) contains details of field locations, data types and any special recognitions criteria. A *source* page that is matched to the fingerprint picks up these parameters.

## Child Object

An object nested within an object of the same type on the Document Hierarchy. In the default *Invoicess* application, as an example, *LINEITEM* fields are child **Field** objects of the parent *DETAILS Field* object. *ItemID*, in turn, is a child of the *LINEITEM* field. *See also: Document Hierarchy.*

## Client/Server Mode

A networking configuration that places a supervisory Taskmaster Server between the Taskmaster Clients and the File Server that holds an application's files. Taskmaster Server is a service that manages the network as it continuously accesses products and applications on the File Server. The *client/server* mode is ideal for configurations employing numerous workstations. *See also: Serverless Mode.*

## Confidence Level

A value between "0" and "9" stipulating a minimum degree of certainty on the part of a Recognition task that it has accurately recognized data in a particular field of a *source* page, or has accurately recognized the page itself. The Required Confidence Levels for a page and its fields are defined in the definition of the task. "0" is the least stringent rating, while "9" means that the task will accept a field's value only if it has interpreted the value with complete confidence. *See also: Field, RuleRunner Task, Source Page.*

## Object

An element of the *Taskmaster* Document Hierarchy. The hierarchy has objects at four levels: **Batch**, **Document**, **Page** and **Field**. An object at any level can be a **parent**, with **child** objects of the same type nested within. *See also: Document Hierarchy.*

## Custom Rule

A rule defined to override or supplement a Global Rule that applies to a particular object of the Document Hierarchy. A custom rule is an attribute of a *fingerprint*. *See also: Global Rule, Fingerprint.*

## Data Source Name (DSN)

Specifications defining the link between a specific data *source*—usually a database such as a *Taskmaster* application's Engine, Admin or Rules database—and an **Open Database Connectivity (ODBC)** "driver." *See also: Application, Engine Database, Admin Database, Rules Database*

In *Taskmaster 6, Release 2* Connection Strings replace DSN's.

## Data Edit Field

An interactive field in a Verification task's *Data Entry Panel*. This field displays the *captured* value for a **Field** object of a page; the accompanying Snippet shows the *image*

of the field and its value. After comparing the two values, a Data Entry operator can make appropriate changes to the value in the data edit field. *See also: Snippet, Data Entry Panel, Captured Value.*

### Data Entry Panel

A window which presents the Verification task's operator with information taken from certain fields of the current page.. Typically, the information appears in two formats: a snippet shows the image of the source field and any value it contains; an interactive data edit field displays a Recognition task's interpretation of the field's value...its *captured* value. The panel also allows an operator to review and modify key identifying codes such as Vendor IDs and Vendor Names, and to set aside pages which RuleRunner was not able to match to a fingerprint in the Fingerprint Inventory.

### Data File (.xml)

A file containing a page's *captured* values. After a Recognition tasks fully processes a *source* page, the task adds these values to a Data file it assembles just for that page. If the succeeding Verification task alters a *captured* value, the task updates the Data file. *See also: Captured Value, RuleRunner*

### Document

An object on the second level of an application's Document Hierarchy. A **Document** object consists of one or more **Page** objects one level below, and is the primary organizing element of the hierarchy's top-level **Batch** object. Batch Contents files (.xml) generated by each task identify the documents in the current batch, and the pages in each document. *See also: Document Hierarchy, Batch, Page, Page Type, Batch Contents File.*

### Document Discrepancy

A procedure that specifies the *expected* number of documents in a batch; counts the documents in each batch after processing by the RuleRunner and Verification tasks; and reports any discrepancies.

### Document Hierarchy

An application's principal organizing structure. The Document Hierarchy has four levels: **Batch, Document, Page** and **Field** – and objects at each level have unique properties. *See also: RuleSet Hierarchy, Workflow Hierarchy.*

### Document Integrity

Settings of the Document Hierarchy that determine the makeup of the documents in a batch, in terms of the *number* and *type* of pages a document must contain. *See also: Document, Page, Document Hierarchy.*

### Document User Data

Additional, operator-specified information about a document. This data appears in the DD line that follows each D line in a task's Batch Contents file. *See also: Batch Contents file, Document Data.*

### Engine Database

The *Taskmaster* database containing processing results according to task, batch, document and page. *See also: Admin Database.*

### Export File

A file (.txt) generated by an Export task at the end of a Main job's workflow. The file consists of Export records containing verified data from the pages of a fully processed batch. The format and content of the Export record is determined by the Export rules governing the extraction and formatting of *captured* values from a page's Data file, and the possible addition of data from other sources. *See also: Captured Values, Export Task.*

### Export Task

The closing task in the Main job of the *default Taskmaster* workflow. In response to Export and ExportClose rules, the task extracts and formats verified data in the Data file of a page and adds it to an Export database or ASCII text file.

### Field

An object at the lowest level of a *Taskmaster* application's Document Hierarchy –and an element of a **Page** object one level above. All fields on a *source* page contain values such as titles; many contain values such as IDs, dates or amounts entered by a form-filler.

*See also: Document Hierarchy, Page, Form-filler*

### File Server

The server on which the files of a **Datacap** configuration reside. Phase 1 of the Installation process copies these files from the Installation CD-ROM to the File Server.

### Fingerprint

The electronic template of a *source* page that can be processed by tasks of the *Taskmaster* workflow. A fingerprint starts off as a scanned page. When a Recognition task is unable to match this page to an existing template in the application's Fingerprint Inventory, an Administrative specialist uses *Rule Manager* technology to establish a fingerprint for the page and for all others which come after it. Once complete, the fingerprint contains a complete set of rules that locate the fields on a corresponding page; read, interpret and validate their data; and export the data to files or databases. A fingerprint consists of an Image file (.tif) and a Fingerprint file (.cco); both are in the application's **Fingerprint** directory. *See also, Rule Manager.*

### Form

A blank, pre-printed entity with fields that are filled in by individuals or organizations (form-fillers) who enter hand-written or machine printed values. *See also: Document, Form-filler.*

### Form-Filler

The individual or mechanism responsible for entering *fluid* values in the fields of a *source* page (such as the *Taskmaster Invoices* page.) A Recognition r task attempts to read and interpret these hand-printed or machine-generated values; the Verification task confirms the accuracy of these interpretations. *See also: Fluid Value*

### Global Rules

A set of default rules in each RuleSet Type that apply to specific objects of the Document Hierarchy.

### High Confidence Value

A value in a field of a scanned page which a Recognition task interprets with a degree of confidence that equals or exceeds the application's requirements. *See also: FormSpec.*

### Image File

An electronic representation a scanned page. The Image file is a digital picture of the actual page, usually stored as a .tif file.

### Job

A set of interrelated *Taskmaster* tasks that operate together to accomplish a specific goal. In most cases, the purpose of a job is to convert paper documents to captured data. A job is object of a workflow. *See also: Workflow Hierarchy, Job, Task.*

### Job/Task Combination

The workflow's *fundamental processing unit*, forged by the assignment of a specific task to a specific job. The workflow prototype of a *Taskmaster* application consists of five tasks assigned to a *Taskmaster* Main job: **Taskmaster Main.Scan**, **Taskmaster Main.ImageFix**, **Taskmaster Main.RuleRunner**, **Taskmaster Main.Verify** and **Taskmaster Main.Export**. *See also: Workflow, Job, Task*

### Job-Task Shortcut

An icon in *Taskmaster's Operations* window that initiates one or more Job/Task combinations. A shortcut is an element of Application Security: it is available *only* to authorized operators, Supervisors and Administrators working from a station with comparable Job-Task permissions. *See also: Job-Task Permissions.*

### Job/Task Permissions

Authorizations permitting individuals to carry out an application's Job/Task combinations. An application's Security Program includes these authorizations in its User, User Group, Station and Shortcut Definitions. *See also: Job/Task Combination, Administrative Privileges.*

### Keyword

A title, phrase or other *alphanumeric* combination which is reasonably likely to appear on a *source* page.. Related keywords are stored in a keyword file: **invnum.key**, for example, lists the many variations of "Invoices Number."

Certain actions of a **Locate** rule attempt to match values in a keyword list with keywords in the current page. Subsequent actions locate the target word that is associated with the keyword and retrieve its *recognized* value. *See also: Target Word*

### Log File

A file that tracks a task's progress as it processes a batch.

### Lookup Database

The database used by Validation rules of the RuleRunner and Verification task to confirm the *fluid* values in certain fields of the current page. *See also: Fluid Value.*

### Low Confidence Field

A field containing data that the RuleRunner task is uncertain about.

### Low Confidence Page

A page that the RuleRunner task determines has one or more Low Confidence fields. A Low Confidence field contains one or more characters that are below the application's recognition threshold

### Main Job

A workflow's primary job, usually consisting of tasks in five categories: Scan, ImageFix, RuleRunner, Verification and Export.

### OCR Field

An Optical Character Recognition field containing typed or computer-printed values.

### OMR Field

An Optical Mark Recognition field containing one or more check boxes to be completed (or ignored) by a form-filler. *See also: Form-filler.*

### Operations Window

A secondary window within the *Taskmaster Window*. The *Operations* window displays shortcut icons an authorized operator, Supervisor or Administrator uses to initiate Job-Task activity. *See also: Job-Task Shortcuts.*

### Operator

An individual authorized by an application's Administrator to run one more tasks: the Scan and Verification tasks usually require direct operator participation. Application Security procedures require a unique User ID and Password for each operator, and formal permission to carry out a specific Job/Task combination. *See also: Job/Task Combinations, Job-Task Permissions.*

### Page

An object on the third level of an application's Document Hierarchy. A **Page** object is an element of a **Document** object one level above, and contains **Field** objects one level below. In the early stages of a Main job, a page is represented by an Image file (.tif). Later, a page is represented by the Image file and a current page that includes a Data file (.xml) *See also: Document Hierarchy, Image File, Data File, Main Job.*

### Batch File

A file generated by a task as it processes a batch. The Page file (.xml) contains identifying codes and processing statistics about the batch, and about the documents and pages the batch contains. After a task has finished with a batch and compiled the file, the next task in the workflow refers to it for processing benchmarks. *See also: Data File.*

### Page Type

The unique identifying property of a **Page** object of the Document Hierarchy. A **Document** object must include one or more **Page** objects – each with its own Page Type value.

### Page Data

A Batch Contents file's standard information about a processed page in a batch. The Batch Contents file's P line contains this data. *See also: Batch Contents File.*

### Page User Data

A Batch Contents file's optional information about a processed page in a batch. Settings in the Scan Task Definition determine the content and format of the data, which appears in the PD line of the Batch Contents file. *See also: Batch Contents File.*

### Panel

*See Data Entry Panel.*



## Privileges

*See Administrative Privileges.*

## Recognition Engine

Software that a Recognition task uses to read and interpret the values in the fields of a page. A *Taskmaster* application uses different engines to read machine-printed and hand-printed text, and to “vote” on values in certain fields.

## Report Viewer

A *Taskmaster* utility that retrieves, organizes and displays an application’s task statistics.

These statistics track the performance of a task as it assembles and processes the contents of a batch – moving the batch from one task to the next until the data on every page has been read and recognized, verified and exported.

*Report Viewer* can present results electronically or on paper. Coverage can be complete, or “filtered” according to parameters you provide. You can call upon *Report Viewer* for spur-of-the-moment details, or for a stream of information generated according to a schedule you define.

## Rescan

An option that a Scan task or FixUp task provides to an operator after a batch has been processed. The procedure re-processes the image of a hard-to-read page. *See also: Scan Task.*

## Rule

A set of instructions that determines how a task is to carry out a narrowly-defined activity involving a single object of the Document Hierarchy.

A rule consists of one or more actions.

Rules are categorized according to their RuleSet Type: **Locate** rules, for example, find fields and their values; **Export** rules extract values from a Data file and add them to an Export file or database. Multiple rules of the same type that present alternative ways to deal with the same object form a RuleSet. *See also: Action, RuleSet, RuleSet Type.*

## Rule Manager

Development technology responsible for definitions of an application’s Document Hierarchy, Global Template, *fingerprints*, rules, RuleSets and RuleSet Types. *Rule Manager* activities take place within the **Rule Manager Window**.

### RuleRunner Task

A background task that responds to pre-defined rules applied to objects of an application's Document Hierarchy.

### Rules-based Application

A **Datacap** application with tasks that operate in response to rules prepared exclusively for the application. These rules are applied directly to objects of a Document Hierarchy that has also been constructed exclusively for the application.

The Document Hierarchy has objects at four levels: **Batch, Document, Page** and **Field**. Rules are categorized according to RuleSet Type: Pre-Recognition, Recognition, Create Documents, Locate, Clean, Filter, Validation, Export and ExportClose.

A *rules-based* application's RuleRunner, Verification and Export tasks run according to rules in one or more categories. *See also: Script-based Application*

### Rules Database

The database with complete information about an application's rules, actions and RuleSets – and their association with fingerprints and individual objects of the Document Hierarchy. *See also: Actions, Rules, RuleSets, Fingerprint, Document Hierarchy.*

### RuleSet

A group of rules of the same RuleSet Type that apply to same object of the Document Hierarchy. As an example, the **Locate** rules that look for the **Date** field on a page belong to the same RuleSet. *See also: RuleSet Type.*

### RuleSet Hierarchy

The internal structure of a process that defines and applies the rules that govern a task's operations. A RuleSet Hierarchy consist of a RuleSet Type, a rule, and the actions that carry out the rule's operations.

### RuleSet Type

A category that defines the purpose of a particular RuleSet. *Taskmaster* specifies nine *default* RuleSet Types: Pre-Recognition, Recognition, CreateDocuments, Locate, Filter, Clean, Validate, Export and Export Close. *See also: Rule, RuleSet*

### Script-based Application

A **Datacap** application that employs standard tasks to process documents in response to the specifications of a FormSpec. Although the set up of individual tasks permits limited customization, an application's versatility results primarily from scripts that expand (or limit) a task's role as it processes a particular page, or a field within a page. *See also: Rules-based Application.*

### Security

Procedures that define and govern the access of individuals and workstations to an application's databases, jobs and tasks; and limit an individual's involvement in administrative activities. *See also: Administrative Privileges, Job-Task Permissions.*

### Serverless Mode

The networking mode connecting Taskmaster Clients directly to the File Server that holds an application's resources instead of through a Taskmaster Server. *See also: Client/Server Mode.*

### Settings File (.ini)

A file whose specifications govern certain aspects of the ways in which a task works, or the content of the *Rule Manager Window*.

### Shortcut Icon

An icon in the secondary *Operations* window of the *Taskmaster Window*; double-clicking on the icon initiates the activities of a Job/Task combination that selects and processes a batch. An application's Security Program assigns properties to an icon and determines who can use the icon. *See also: Job-Task Shortcuts.*

### Snippet

A field of a *Taskmaster Data Entry Panel* that shows the image of a field in a page, and of any data it contains. A snippet is paired with an interactive data edit field; this field contains the RuleRunner task's interpretation of the field's value. After comparing the values in each field, a Verification task operator can make any necessary changes. *See also: Data Edit Field, Data Entry Panel.*

### Source Page

A page represented by the **Page** object of the Document Hierarchy that has **Field** objects. Rules associated with the **Field** objects of a *source* page determined how the workflow recognizes, extracts, verifies and validates the fields' values – and exports them to a file or database. *See also: Page, Document Hierarchy, Static Value, Fluid Value.*

### StartBatch Data Panel

A panel that appears when a KScan task begins: the panel asks the operator for specifics about the contents of the new batch. The panel's **Pages in Batch** amount becomes the workflow's *expected* number of pages in the batch.

### Station

A computer running Taskmaster Client software and authorized to carry out one or more Job/Task combinations. Also known as workstations, *Taskmaster* security assigns each an

access code; an operator can use this computer to run a task *only* if it has a valid Station ID (and the operator has a valid User ID and Password.) See also: Security.

### Station Monitor

An administrative window with details about each workstation in an application's *Taskmaster* network.

### Super Snippet

A type of snippet. The difference between a regular snippet and a super snippet is that a Verification task's operator can view a super snippet for a data entry field even though the panel does not contain a snippet for that field. *See also: Snippet, Data Edit Field*

### Target Directory

A folder on the File Server containing the **Datacap** files that have been copied from the Installation CD-ROM during the opening phase of the Installation process. Installation's second phase, this directory becomes the unambiguous target to which Taskmaster Server will point as it searches for applications and data, if *Taskmaster* is running in a *client/server* mode. Similarly, Stage 3 will provide any client running in a *serverless* mode with a straight path to the target files. *See also: Client/Server Mode, Serverless Mode, File Server.*

### Task

The processing component of a *Taskmaster* Workflow Hierarchy. *Taskmaster* includes tasks in five *default* categories: KScan, ImageFix, RuleRunner, Verification and Export. A task must be part of a job; a job must be part of a workflow. The task's identity is defined in *Taskmaster* when it added to the Workflow Hierarchy. However, because a *Taskmaster* task is a product of the *Batch Pilot* workshop, its setup specifications and detailed settings are elements of the underlying Task Project. *See also: Workflow Hierarchy, Batch Pilot.*

### Taskmaster

The administrative engine of the *Taskmaster* application. *Taskmaster* defines and manages the workflow—beginning with the physical scanning of a page, and ending when the values in the fields of a page have been verified and exported.

### Taskmaster Administrator

A secondary window of the *Taskmaster Window*, used to define an application's Workflow Hierarchy and Security Program.

## Taskmaster Client

The object of the *Taskmaster* configuration responsible for initiating and monitoring the Main job's tasks. Taskmaster Clients access the application's databases and supporting files through Taskmaster Server. The application's Security Program defines the scope of a client's activity: Administrative clients can carry out a full range of tasks and supplementary procedures. *See also: Taskmaster Server, Databases.*

## Taskmaster Web

A processing environment that uses a *Taskmaster* Web Site to link remote Taskmaster Clients to an application's administrative and operating functions.

## Taskmaster Server

The component of a *Taskmaster* configuration responsible for linking the system's Engine and Admin database with Taskmaster Clients *if* *Taskmaster* has been set up in a client/server mode. *See also: Taskmaster Server, Databases, Client/Server Mode.*

## Taskmaster Server Service

The *service* that carries out procedures otherwise assigned to Taskmaster Server.

## Task Monitor

A secondary, display-only window that appears as soon as a task selects a batch for processing, then provides up-to-the-second data as it tracks the task's progress.

## .tif File

*See Image File.*

## User

1. An individual who fills out the fields of a paper copy of a form.
2. An individual who has access to one or more objects of a *Taskmaster* application and its databases. The application's Security Program provides each user with a User ID and Password, and specify the limits of a user's processing and database access "privileges." *See also: Form-Filler, Security.*

## Validation

The process of making sure that a field's value meets specified criteria. For example, the criteria for a field might be that the field can only contain a number within a range of numbers. The process is governed by **Validate** rules that apply to specific **Field** objects of the Document Hierarchy, and are called by a Main job's RuleRunner and Verification tasks.

### Verification Task

A core task of the *Taskmaster* Main job. This operator-intensive task retrieves information from the Image (.tif) and Data (.xml) files of a “problem” page and displays it in the fields of the application’s **Data Entry Panel**. After review by the operator, the task moves to the next problem page until the batch is verified. *See also: Data Entry Panel.*

### Voting

Internal procedures of a RuleRunner task that use multiple Recognition engines to read and interpret the value in a field. These procedures include an algorithm that settles on the most satisfactory interpretation of the value. *See also: Recognition Engine*

### Workflow

The top level of an application Workflow Hierarchy. A job, the tasks assigned to the job, any child jobs that branch from the job, and the tasks of these child jobs. *See also: Job, Task, Workflow Hierarchy.*

### Workflow Hierarchy

The formal structure that a *Taskmaster* application employs to retrieve, interpret, verify and store the data on a scanned page. A Workflow Hierarchy has three tiers. At the top is the workflow itself, which focuses on a particular form or operates within a particular environment. Jobs are objects of a workflow and occupy the second tier: a workflow usually has a Main Job, and may have child jobs to handle special functions. Tasks are objects of a job, and are on the hierarchy’s third tier. *See also: Job, Task, Workflow.*

### Zone

A tightly-defined area that surrounds a **Field** object of a *fingerprint*. The area’s parameters – its size and location – are properties of the object, and are transmitted to pages that rely on this *fingerprint* for guidance.