**C868 – Software Capstone Project Summary**

**Task 2 – Section C: User Guide**



|  |  |
| --- | --- |
| **Capstone Proposal Project Name:** | FS InfoCat – File System Information Cataloging Tool |
| **Student Name:** | Leonard T. Erwine |

Table of Contents

[Overview 2](#_Toc83611418)

[Key Concepts 3](#_Toc83611419)

[Crawl 3](#_Toc83611420)

[Volume 3](#_Toc83611421)

[Hash Code 3](#_Toc83611422)

[Redundancy Set 3](#_Toc83611423)

[File System 3](#_Toc83611424)

[Symbolic Name 3](#_Toc83611425)

[The Start Screen 3](#_Toc83611426)

[Defining Crawl Parameters 4](#_Toc83611427)

[Crawl Configuration Edit Page 4](#_Toc83611428)

[Display Name 5](#_Toc83611429)

[Starting Subdirectory 5](#_Toc83611430)

[Current Status 5](#_Toc83611431)

[Next Scheduled Start 5](#_Toc83611432)

[Auto reschedule / Manual only 5](#_Toc83611433)

[Limit Duration 6](#_Toc83611434)

[Max Recursion Depth 6](#_Toc83611435)

[Limit Total Items 6](#_Toc83611436)

[Command buttons on lower right corner 6](#_Toc83611437)

# Overview

FS InfoCat is an application which crawls through subdirectories and collects information about files. This information is stored in a local database where it can be sorted analyzed. The main goal of the application is to assist with locating possible duplicate files and as a tool to simultaneously view the contents of all subdirectories that have been controlled.

## Key Concepts

### Crawl

This refers to the process of scanning through nested subdirectories to collect information about files.

### Volume

Synonymous with other terms like “Drive”, “File Share” and “Logical Volume”, it is a logical folder structure containing zero or more files.

### Hash Code

Sometimes called an MD5 Hash Code, this is a cryptographically generated sequence of byte values that can indicate whether 2 files may be identical.

### Redundancy Set

A grouping of files which share the same size and hash code.

### File System

From the perspective of this app, it is a database record which describes a files system type.

### Symbolic Name

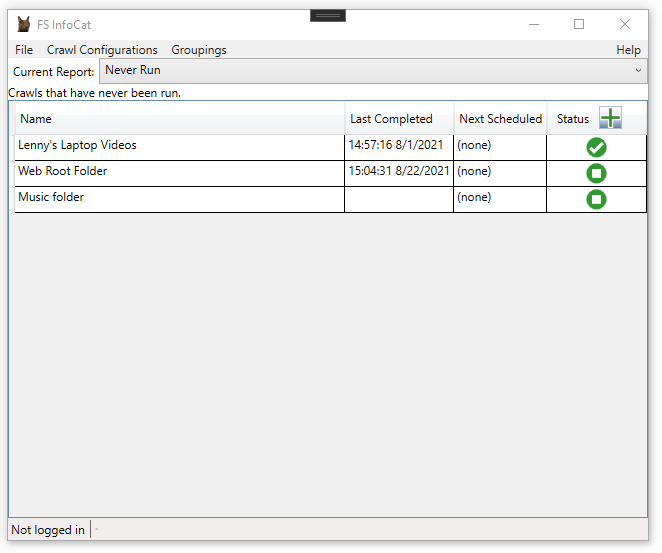
An abbreviate name for a file system type.

## The Start Screen

When the application is opened, you will typically recent and upcoming crawl job report.

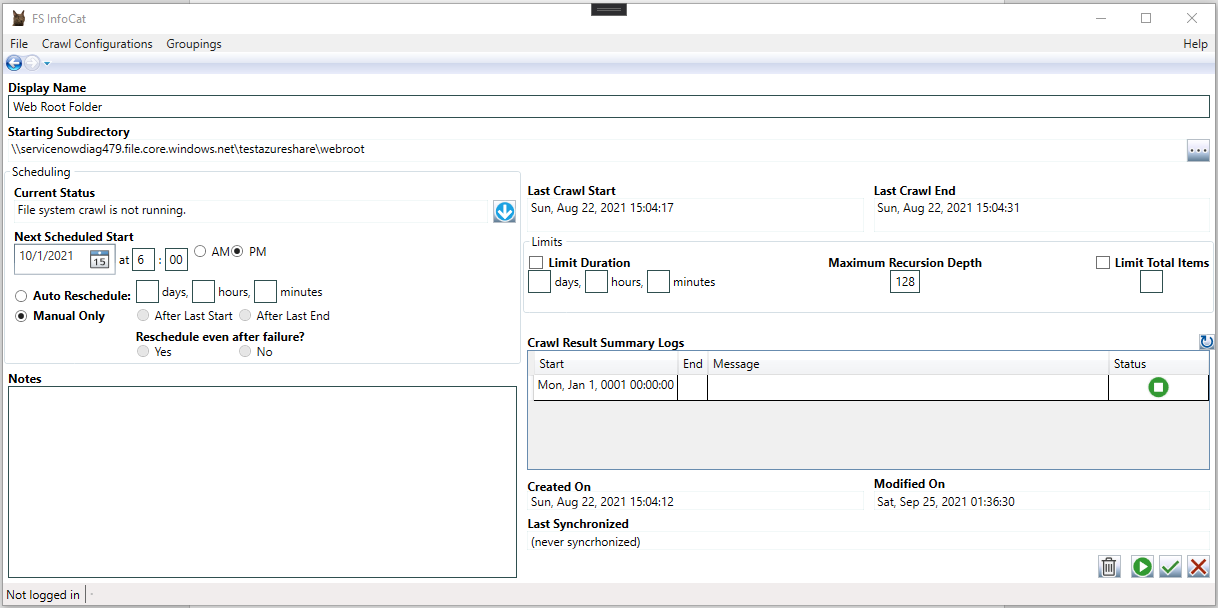
# Defining Crawl Parameters

Before any crawls can take place, you must create a Crawl Configuration record, which represents specific crawl parameters, such as how long it can run or how far it can recurse into subdirectories. To add a new Crawl Configuration, you can click the Add button toward the upper-right of a crawl configuration listing page, or you can select “Crawl Configurations” and then “New Crawl”.



## Crawl Configuration Edit Page

Following is an example Crawl Configuration edit form.



### Display Name

This is the display name to show for this crawl configuration.

### Starting Subdirectory

The starting subdirectory to be scanned. From this location, the crawler will recurse into nested subdirectories.

### Current Status

This is a message indicating the status of the current crawl. This also has a graphic to the right, which gives a quick visual of the status.

### Next Scheduled Start

This is the date and time when the next crawl is supposed to begin.

### Auto reschedule / Manual only

Indicates whether the configured crawl will be rescheduled after completion.

### Limit Duration

Use this to limit how long the crawl can run.

### Max Recursion Depth

Use this to limit how many levels deep it can crawl through subdirectories.

### Limit Total Items

This limits the total number of items processed.

### Command buttons on lower right corner

Clicking the leftmost button will save your changes and immediately start a new crawl.

The middle button will save your changes and return you to the previous page.

The rightmost button will discard your changes and return you to the previous page.