



This repository Search

Explore Gist Blog Help



terrywbrady



Georgetown-University-Libraries / File-Analyzer

forked from terrywbrady/File-Analyzer

Unwatch 5

★ Unstar 13

Fork 7

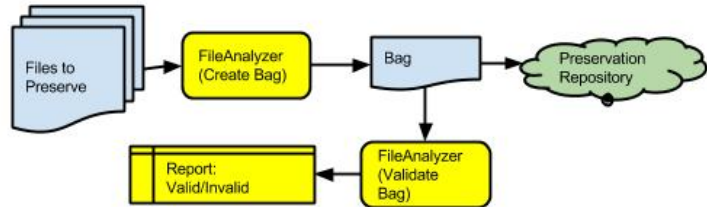
# Bagit automation

Terry Brady edited this page 4 days ago · 6 revisions

Edit

New Page

This rule was written to provide **BagIt** functionality through the File Analyzer. This rule was designed for users who perform other digitization tasks using the File Analyzer.



## Verify Bag

Given an existing Bag, verify that the integrity of the Bag contents.

Criteria Import Progress Details Merge

Root Directory to Scan.

C:\Temp\bags ... Recent

Action to perform on Files

Verify Bag Max Items: 1000000

File Test Desc File Filter Criteria File Test Properties Advanced Options

Description of test to be performed

This rule will scan for directories with names ending with '\_bag'. BagIt verifyvalid will be run on the bag.

Criteria

Import

Progress

Details

Merge

Bag1

Bag2

Verify Bag for .../Temp/bags

Remove Tab

Bag Status

Summary Counts

Bag Path	Bag Status	Item Count	Message
Marc_bag	VALID	83	
lone_bag	INVALID	27	(error) Payload manifest manifest-md5.txt contains invalid file(s): [data/
lone_bag_bag	INVALID	0	(error) Payload file data/lone_bag/data/ingest/folder_00002/EmbargoC...
lone_bag_bag\data\lone_bag	INVALID	27	(error) Payload manifest manifest-md5.txt contains invalid file(s): [data/
two_bag	VALID	27	

## Create a Bag from an existing directory

Pages 46

- Home
- File Analyzer Component Packages
- Installation instructions
- File Analyzer Stories
- File Analyzer Use Cases at Georgetown University
- Latest Features
- User Interface Overview
- Command Line Interface
- Batch Processing
- Coding new File Test Rules and new File Import Rules

- File-Analyzer-Training-Code4Lib-2015

Clone this wiki locally

<https://github.com/Georgetown-University-Libraries/File-Analyzer/wiki/Bagit-automation>

Clone in Desktop

Criteria Import Progress Details Merge

Root Directory to Scan.

C:\Temp\Marc

Action to perform on Files

Create Bag  Max Items: 1000000

File Test Desc File Filter Criteria File Test Properties Advanced Options

Description of test to be performed

*This rule will bag the contents of the selected directory 'X' and create a new bag directory name 'X\_bag'.  
Some thought is still needed on how to iteratively bag up subfolders or files into independent bags.*

File Analyzer

Criteria Import Progress Details Merge Bag1

Create Bag for .../Temp/Marc

Bag Status

Summary Counts

Source	Bag	Bag Status	Item Count	Message
Marc	C:\Temp\Marc_bag	VALID	83	

## TODO's

### Enable the source of a bag to be any of the following options

- Root folder
- Subfolders of root folder
- Files contained within root folder (matching by name)

Assumption: all 3 are needed. Currently only the root folder is supported.

### Enable the creation of directory bags and zip file bags

Ensure that the validation task works in both instances.

Assumption: directories. Zip support would be useful to add

Add a custom footer

