View on GitHub

File-analyzer

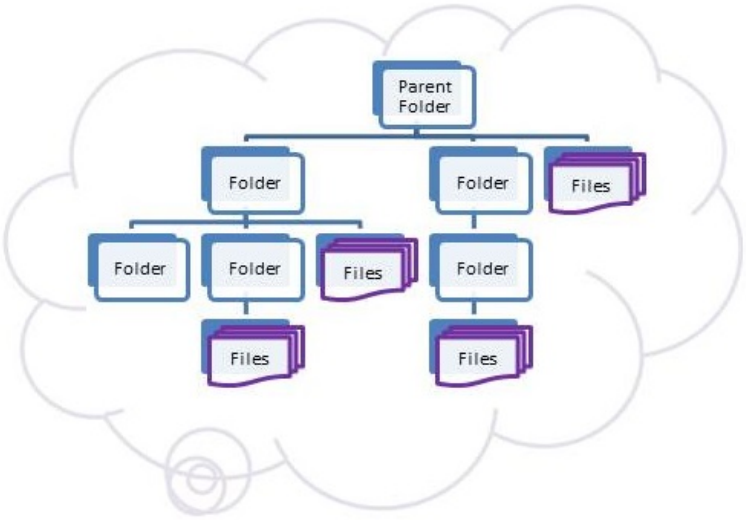
File Analyzer and Metadata Harvester

Download this project as a .zip file Download this project as a tar.gz file

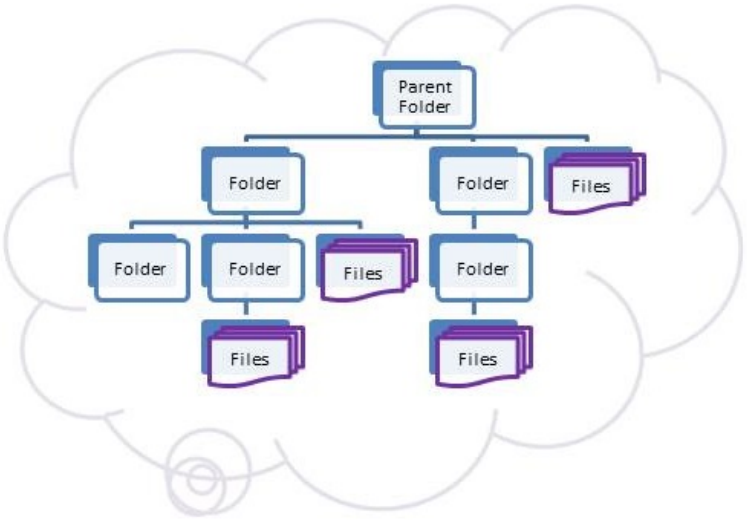
PURPOSE The File Analyzer is a general purpose desktop (and command line) tool designed to automate simple, file-based operations. The File Analyzer assembles a toolkit of tasks a user can perform.

The tasks that have been written into the File Analyzer code base have been optimized for use by libraries, archives, and other cultural heritage institutions.

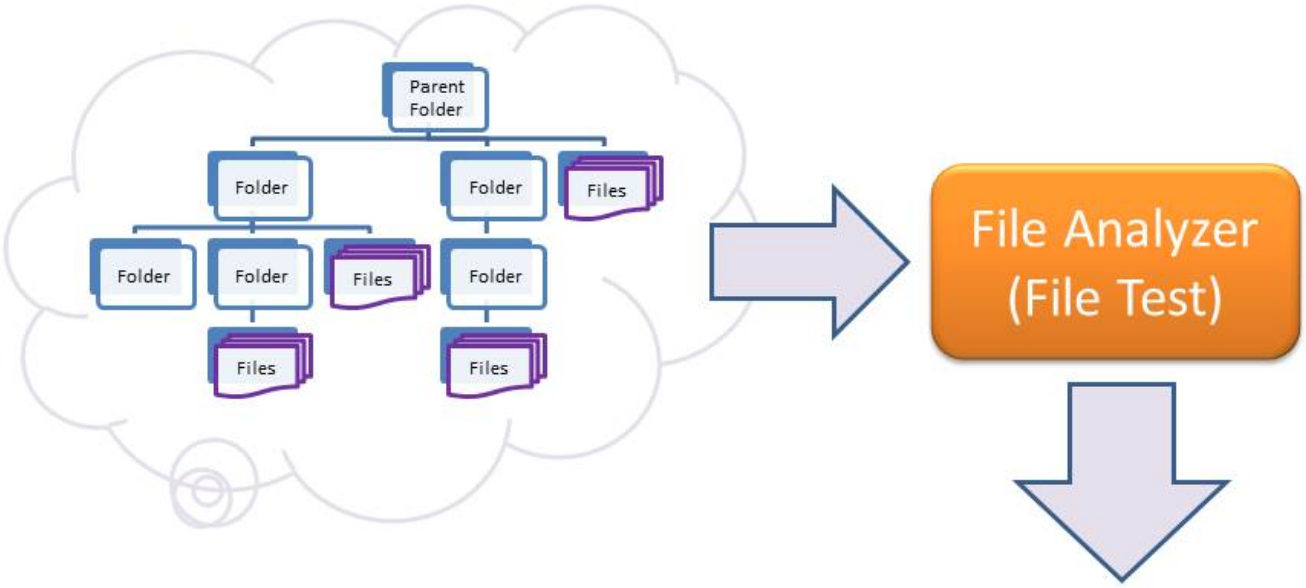
See the File Analyzer Wiki for detailed instructions.



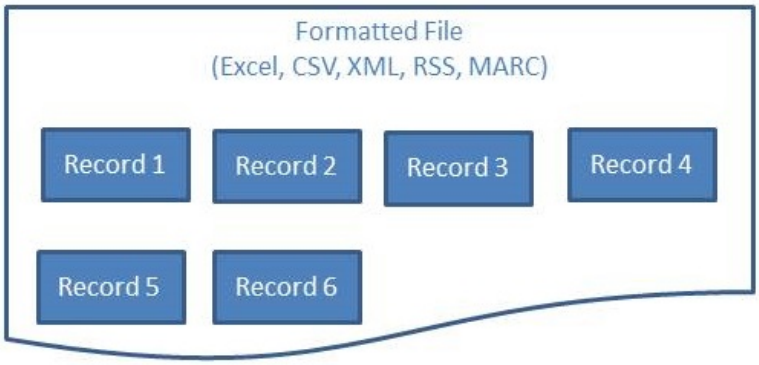
You have a collection of files and you need to do something with those files...



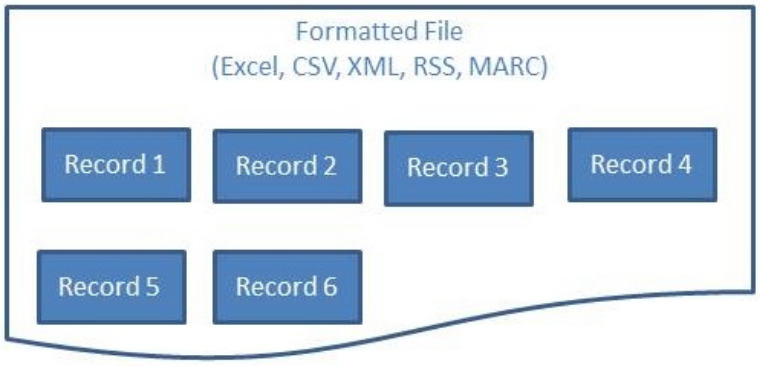
1. Identify which files meet a specific criteria: Pass/Fail
2. Extract information about each file
3. Convert each file into a new format



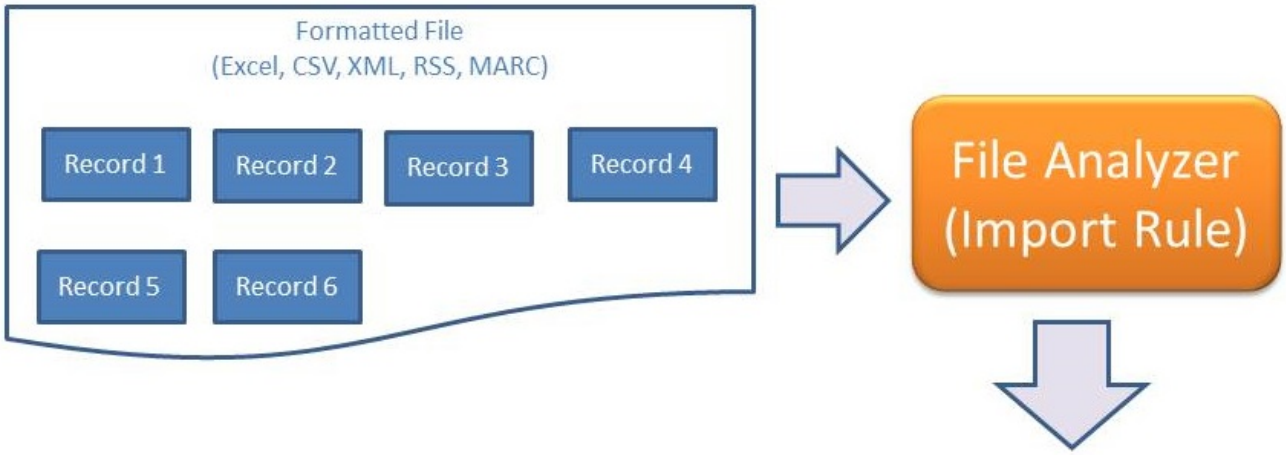
File Name	Status	Author	Title	Size	Message
File1.pdf	PASS	Author 1	Title 1	1.2 MB	
File2.pdf	FAIL		Title 2	2.3 MB	No Author
File3.pdf	PASS	Author 3	Title 3	1.7 MB	
File4.pdf	WARN	Author 4	Title 4	11.2 MB	Large File



You have a formatted file containing of several records and you need to do something with each of those records...

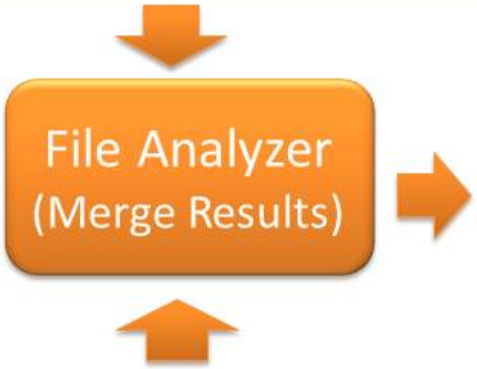


1. Identify records that meet a certain criteria
2. Extract information about each record
3. Generate a new file from the individual records



Record	Status	Author	Title	Message
00011	PASS	Author 1	Title 1	
00014	FAIL		Title 2	No Author
00120	PASS	Author 3	Title 3	
00121	WARN	Author 4A	Title 4	Large File

File Name	Status	Author	Title	Size	Message
File1.pdf	PASS	Author 1	Title 1	1.2 MB	
File2.pdf	FAIL		Title 2	2.3 MB	No Author
File3.pdf	PASS	Author 3	Title 3	1.7 MB	
File4.pdf	WARN	Author 4	Title 4	11.2 MB	Large File



Title	Author (left)	Author (right)	Status
Title 1	Author 1	Author 1	Matches
Title 2			Matches
Title 3	Author 3	Author 3	Matches
Title 4	Author 4	Author 4A	Mismatch

Record	Status	Author	Title	Message
00011	PASS	Author 1	Title 1	
00014	FAIL		Title 2	No Author
00120	PASS	Author 3	Title 3	
00121	WARN	Author 4A	Title 4	Large File

Read File Analyzer Stories to learn more about the problems that have been solved with the application.

See the File Analyzer Wiki for detailed instructions.

Please view the project README for license information and installation requirements.

File-analyzer maintained by Georgetown-University-Libraries

Published with GitHub Pages