Working with Regular Expressions (Regex)

Introduction

This guide explains how to use Regular Expressions (REGEX) in the ADF applications to build advanced search patterns for files, folders, and text content. Regular expressions provide a flexible way to match complex patterns such as email addresses, dates, URLs, and code fragments.

Regular expressions can be used in keyword Captures, in the targeted and excluded folders of File Captures, and in the header byte sequences of file types. Review Scan Results shows the matches of regular expressions, but its search fields do not accept regular expressions.

This document applies to the following applications		
ADFPRO Mobile & Computer Triage	DEI Digital Evidence Investigator	MDI Mobile Device Investigator

When Regular Expressions Can Be Used

Regular expressions can be used in the following areas of the ADF desktop applications:

Area	Screen or Function	Description
Keyword Captures	Scan Setup & Key Management > Manage Captures > New Capture > Search for keywords	In the Search Expressions table, set the Search Type of the expression to Regular expression. The Selected Search Type list above the table sets the type of the rows you add; APPLY sets it for the checked rows.
Keywords Added Before a Scan	Scan screen of a live or boot scan from a Collection Key, or of a Targeted Extraction Only scan (Investigate Device) > Add Keywords	The same keyword table and Search Type list as in a keyword Capture, including Regular expression.
Keyword List Imports	Define Keyword Capture > Import List	In a UTF-8 CSV file with a header row, the value 4 in the “search type” column marks a regular expression. Enclose an expression that contains a comma, for example `[0-9]{1,3}`, in double quotes, and double any quote inside it.

Copyright ©2016-2026 ADF Solutions, Inc. All rights reserved. U.S. Pat. No. 8,219,588 and 7,941,386. ADF PRO, Digital Evidence Investigator and Mobile Device Investigator are trademarks or registered trademarks of ADF Solutions, Inc. All other trademarks referenced herein are the property of their respective owners.

File Type Definitions	Define Files to Collect, Define Files to Search or Define Files to Hash > File Types > Specific Files > View > Add File Type	Only the Header Byte Sequence column accepts a regular expression, limited to bytes written as `\xHH`, `.` for one byte and `.{n}` for n bytes, for example `\x12\x4e.{2}\x2b` (up to 100 characters). The sequence must match from the first byte of the file. A dot does not match the byte 0x0A; write that byte as `\x0a`. The sequence is checked only by Thorough identification for all files, and by Thorough identification for files without extensions for files that have no extension. The Extension and File Name columns do not accept regular expressions. Changes to file types affect all Captures.
Excluded and Targeted Folders	Define Files to Collect, Define Files to Search or Define Files to Hash > File Sources > Targeted folders > View > Include Folder; the same editors > Options > Excluded folders > View > Exclude Folder	Each entry is a regular expression that must match the whole path of a file, from the root of its file system (without a drive letter), with / between folders. Letter case is ignored. For example, `/users/.*` covers all files under the Users folder, and `.*/desktop/.*` all files in any Desktop folder. The lists are shared by all Captures; each Capture uses the entries checked in it.
Document Viewer Match Sidebar	Review Scan Results > a keyword Capture > select an expression > Details > Preview	For an expression of the Regular expression type, the Preview tab of a document or PDF lists the matches found in its content next to the document; click a match to find it in the document. With no expression selected, the list shows the matches of all expressions. The Excerpts tab shows each match in its context.

Regular expressions are not supported in the following places:

- The Search Tables keyword search of Review Scan Results. It finds words and the beginnings of words, and it rejects most regular expression characters.
- The Search function of the Preview tab in the Details pane. It finds the typed text.
- The Extension and File Name columns of a file type.

Regular Expression Engine

The ADF applications match keyword search expressions and header byte sequences with the RE2 engine developed by Google. RE2 is optimized for speed, safety, and predictability. Its run time grows linearly with the length of the searched text, so an expression cannot stall a scan through “catastrophic backtracking.”

The benefits include:

- Predictable performance on large data sets.

Copyright ©2016-2026 ADF Solutions, Inc. All rights reserved. U.S. Pat. No. 8,219,588 and 7,941,386. ADF PRO, Digital Evidence Investigator and Mobile Device Investigator are trademarks or registered trademarks of ADF Solutions, Inc. All other trademarks referenced herein are the property of their respective owners.

- Supports common syntax familiar to PCRE users.
- Unicode-aware: Unicode classes such as `\p{L}` and the `(?i)` flag work for all languages, but `\b`, `\d`, `\w` and `\s` cover only ASCII characters.
- Does not support constructs that need backtracking, such as lookarounds and backreferences (see Unsupported Syntax).

How a Scan Applies an Expression

- Matching is case-sensitive. Start an expression with `(?i)` to ignore letter case; most expressions supplied with the application use it.
- Each text is searched on its own: the path of a file or folder (see the next point), the content of a file, each metadata value, and each field of an artifact record, such as the subject or the body of an email. `^` and `$` match at the start and the end of such a text. Add `(?m)` to make them match at the start and the end of every line; with Windows line breaks, use `\r?$` for the end of a line.
- In the File and folder names scope, a file or folder is found when the expression matches its name, and the matches are then taken from its full path, for example `…/Users/John/report.docx`, so folder names in the path can match too. `^` therefore does not mark the start of a name; use `(?:^|/)` instead; it also matches at the start of every folder name. `$` marks the end of a name.
- At most 50 matches of each expression are kept for each searched text, and the Scan Log reports “Encountered more than 50 instances of keyword”. Matches can overlap: `(?i)[A-F0-9]{32}` finds 9 matches in a 40-character hexadecimal value. Put `\b` at both ends to match whole values only.
- The application does not check the syntax of an expression when you enter or import it. An expression that RE2 cannot use, for example one with a lookahead, is skipped when the scan starts and finds nothing; the Scan Log does not report it. Test new expressions before you use them in a scan (see Testing Expressions).
- Avoid expressions that can match an empty string, such as `a*` or `(abc)?`. Such an expression can hide the matches of the other expressions in the same Capture.

Supported Syntax

Category	Examples	Description
Character Classes	`[A-Za-z0-9_]`, `[^0-9]`	Matches any character in the set, or any character outside it when the set starts with ^.

Copyright ©2016-2026 ADF Solutions, Inc. All rights reserved. U.S. Pat. No. 8,219,588 and 7,941,386. ADF PRO, Digital Evidence Investigator and Mobile Device Investigator are trademarks or registered trademarks of ADF Solutions, Inc. All other trademarks referenced herein are the property of their respective owners.

Predefined Classes	`\d`, `\w`, `\s`	An ASCII digit; an ASCII letter, digit or underscore; a space, tab, line break or form feed. `\D`, `\W` and `\S` match any other character.
Unicode Classes	`\p{L}`, `\p{Lu}`, `\p{Cyrillic}`	A letter of any language, an uppercase letter, a character of a script.
Any Character	`.`	Any character except a newline (\n).
Quantifiers	`*`, `+`, `?`, `{n}`, `{n,}`, `{n,m}`	Repeat the preceding item: zero or more, one or more, zero or one, exactly n, at least n, or n to m times (at most 1000). Add ? to repeat as few times as possible, for example `*?`.
Anchors	`^`, `$`	Match the start or end of the searched text, or of each line with `(?m)`.
Alternation	`cat|dog`	Matches either alternative.
Grouping	`(abc)`, `(?:abc)`	Groups sub-patterns.
Word Boundaries	`\b`	Matches the edge of a word of ASCII letters, digits and underscores.
Escaping	`\.`	Match literal dot.
Flags	`(?i)`, `(?m)`, `(?s)`	Ignore letter case; let ^ and $ match at every line; let . match a newline. A flag applies from where it stands; `(?i:abc)` applies it to a group only.

Unsupported Syntax

Feature	Example
Lookarounds	`(?=abc)` / `(?!abc)` / `(?<=abc)` / `(?<!abc)`
Backreferences	`\1`, `\2`
Conditional expressions	`(?(1)yes|no)`
Atomic groups and possessive quantifiers	`(?>abc)`, `a*+`

Copyright ©2016-2026 ADF Solutions, Inc. All rights reserved. U.S. Pat. No. 8,219,588 and 7,941,386. ADF PRO, Digital Evidence Investigator and Mobile Device Investigator are trademarks or registered trademarks of ADF Solutions, Inc. All other trademarks referenced herein are the property of their respective owners.

Free-spacing mode	`(?x)`

Example Patterns

Expression	Description
`(?i)\b[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,}\b`	Detects email addresses.
`https?://[^\s]+`	Detects web URLs.
`[0-9]{4}-[0-9]{2}-[0-9]{2}`	Detects dates in format YYYY-MM-DD.
`(?i)password`	Case-insensitive match for “password.”
`(?i)\b[A-F0-9]{32}\b`	Detects MD5 hash strings.

Testing Expressions

The application does not check an expression when you save it. Test new expressions on sample text before you use them in a Capture. In ADF PRO and Digital Evidence Investigator, add the expression to a test Capture, put sample files with text that should and should not match into a folder, and scan that folder (Add Folder on the Scan screen).

The following web resources can also help. Never enter case data into web tools.

Purpose	Resource
Official RE2 syntax reference	https://github.com/google/re2/wiki/Syntax
Interactive regex tester (select the Golang flavor, which uses the RE2 syntax)	https://regex101.com/



Many regex testers default to PCRE. Select the Golang flavor, which uses the RE2 syntax, so that the tester accepts the same constructs as the scan. The scan is case-sensitive unless the expression uses `(?i)`.

Copyright ©2016-2026 ADF Solutions, Inc. All rights reserved. U.S. Pat. No. 8,219,588 and 7,941,386. ADF PRO, Digital Evidence Investigator and Mobile Device Investigator are trademarks or registered trademarks of ADF Solutions, Inc. All other trademarks referenced herein are the property of their respective owners.