IT Audit Test: Keyword Search In Memo Field (Michael Kano)

Keyword Searches
Although individual keywords can easily be detected using the FIND() function, detecting multiple keywords using FIND() can become a tedious and error-prone process. This script will use a master keyword list in the IT_Audit_Keywords.txt file as the basis for the search. The output will consist of the keyword and the record number in the original file in which each keyword was detected. To edit the list, open the IT_Audit_Keywords.txt file and add or delete any keywords as needed, using only lowercase text and only one word per line.

How the Script Works
Once the user has identified the table and the field within that table to be searched, it parses the memo field, breaking it down into its individual words. Each word and the record number in which it occurs are extracted to a new table that will contain one record for each word in the original field. Standard punctuation marks are excluded, and the words are converted to lower case for exact matching as required by the JOIN command.

For example, if the Memo field in a table contains the following two records,

Record	Memo
1	Identify the person responsible.
2	Release the funds.

the resulting table (Wordsave_Data in the script) would look like this:

Record	Keyword	Record number in source file
1	identify	1
2	the	1
3	person	1
4	responsible	1
5	release	2
6	the	2
7	funds	2

This table would then be joined to the IT_Audit_Keywords table, and any keyword matches would be extracted to a new table, Events_With_Keywords.

The script syntax and the detailed explanations of each line are on the following page.

Syntax	Explanation
`ACCEPT "Select the table." FIELDS "xf" TO v_Table_Select`	Prompts user to select table containing the field to be searched. Table named stored in v_Table_Select variable.
`OPEN %v_Table_Select%`	Opens table selected by user.
`ACCEPT "Select the memo field." FIELDS "C" TO v_Field_Select`	Prompts user to select the field to be searched. Field name stored in v_Field_Select variable.
`V_Word_Save = blanks(50)`	Initializes variable for capturing individual words.
`V_Num_Loops = 0`	Initializes variable for counting number of times LOOP will be executed in each record.
`V_Num_Seg = 0`	Initializes variable for the segment (word) count in each record.
`GROUP`	GROUP
`v_Word_Save = blanks(50)`	Initializes variable for capturing individual words.
`v_Num_Seg = 0`	Initializes variable for the word count in each record.
`v_Num_Loops = OCCURS(ALLTRIM(%v_Field_Select%)," ") + 1`	Number of times LOOP will be executed in current record determined by the number of words in the memo field. Blank space between each word used as the basis for count.
`LOOP WHILE v_Num_Loops > 0`	LOOP with condition: Execute as long as remaining words in the current record > 0.
`v_Num_Seg = v_Num_Seg + 1`	Segment incremented by 1.
`v_Word_Save = SPLIT(EXCLUDE(%v_Field_Select%,",.?-")," ",v_Num_Seg)`	SPLIT command used to capture current word in the field.
`EXTRACT LOWER(v_Word_Save) AS "Keyword" RECNO() AS "Record_Number" TO Wordsave_Data`	Captured word in lowercase and record number extracted to Wordsave_Data file.
`v_Num_Loops = v_Num_Loops -1`	Number of LOOP executions remaining reduced by 1.
`END`	End of LOOP.
`END`	End of GROUP.
`OPEN Wordsave_Data`	Open Wordsave_Data file containing extracted words.
`OPEN IT_Audit_Keywords SECONDARY`	Keyords file opened as secondary file for upcoming JOIN.
`JOIN PKEY Keyword FIELDS Keyword Record_Number SKEY Keywords TO "Events_with_Keywords" OPEN PRESORT SECSORT`	Wordsave joined to Audit_Keywords to isolate any keywords detected in the Wordsave_Data file.