

Technical Disclosure Commons

Defensive Publications Series

April 09, 2018

Configurable data import of multiple Excel

Daniel Hoppe

Bertrandt Ingenieurbüro GmbH

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Hoppe, Daniel, "Configurable data import of multiple Excel", Technical Disclosure Commons, (April 09, 2018)
https://www.tdcommons.org/dpubs_series/1139



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

CONFIGURABLE DATA IMPORT OF MULTIPLE EXCEL FILES

Technical task:

The task of the technical innovation is to provide the possibility of creating a summary of several Excel files.

Initial situation:

The popular spreadsheet program Microsoft Excel works with two-dimensional tables that have rows and columns. Within these tables, data is usually stored in a line-oriented manner. The respective attributes are maintained in separate columns and have column headings.

In practice, there may be more files in the same or similar pattern.

Not infrequently, users create files individually by inserting additional columns, for example.

Now a summary of the data from several files is required. There are two files with the following table contents, see figures 1 and 2.

Required is a summary of the table contents with the column headers „Project“, „Project ID“, „2017“, „2018“ and „2019“. For the summary, however, only „current projects“ should be listed.

The desired result would look like this, see Figure 3.

A common way to create such a view in Microsoft Excel would be to format each table into the target format. Afterwards, the

Manually copy table contents into the summary.

This manual procedure for creating a summary from several Excel files is time consuming and not very reliable.

Excel does not provide a suitable function here.

Solution:

An additional functionality for the spreadsheet program imports the desired data from selectable files and generates a summary.

The import is easily configurable and robust in the sense that the import functionality can handle the individuality of the individual files, performs error checks and creates an import protocol.

In addition, the import functionality offers an alternative line-oriented output form, see Figure 4.

The description of the functionality is based on a sample application with the output tables in Figures 1 and 2.

1. Installation of the import regulation

First, the import is configured.

A possible import rule is shown in Figure 4.

2. The column „B“ contains keywords whose meaning is further defined in column „C“ or „Type“. Type „SS“ denotes a column heading named „Project“.

The importer will analyze all cells below for import in this defined column, see Figure 5.

- An import of a line will only take place if there is no empty cell in this column. Only the SS type may be assigned once. Optionally, a filter criterion for the line to be read can be entered in column „G“. The filter criteria that apply in the Microsoft Office environment apply: ? for any Characters, * for any string, # for numbers, and so on. For example, a filter criterion might look like this: „?? ## *“. This would allow the importer to read only lines that contain a keyword in the o.g. Have format. That is, in the first and second place may be any character; in third and fourth place a digit is expected and then a can any expression. A valid name would be „BE77v“.

Type „S“ indicates column headings that the importer should read. In the summary or output, another column heading can be selected in column „D“ or „Output column header“ in the import rule.

The column number in the output is specified in column „F“. That is, the order of the columns in the summary is individually adjustable.

In the column „G“ a data check can be entered. At a Keyword of type „S“, the data to be read can be checked as to whether it is a date or a number. Will not work specified, no check is made.

The result of the test will be in one Import log listed. Optionally, another test specification could be defined here. Sometimes, in output tables, sums are formed by formulas that you do not want to import. In the test specification, Excel formulas to be defined (eg „= SUM“ or „= PARTIAL RESULTS“) could be used by the Reading process can be excluded.

Of particular importance are the keywords of type „KB“ and „KE“. The importer reads in the column of the keyword „SS“ all data below this column header. With the type „KB“ the importer becomes the Start reading process from the specified keyword. Is a Keyword specified with the type „KE“, the importer is only up to Import this keyword. The keywords of type „KB“ and „KE“ are optional. Without their specification the import of all data takes place below the column heading of type „SS“. The column „H“ is only an explanation and has no effect on the import process. The import rule can be saved for later re-use.

4. Selection of the files to be imported

This function consists of a selection dialog of files or complete folders in which the files are located. The files to be imported are successively loaded into a list. That is, for example, the files may be added from a folder or individual files in this list. Possibly. You can also remove files from the import list here.

5. Further import configuration

Here the user is offered that the name of each imported line

The source file and / or the name of the worksheet is listed in the import.

6. Implementation of the import / result

The import is carried out or prepared according to the selected import files, the import rule and the further configuration.

An Excel file can consist of several tables. The importer „reads“ every visible table. If the column heading of type „SS“ is not on the Table sheet is found, the sheet is not imported.

6. Implementation of the import / result

The import is carried out or prepared according to the selected import files, the import rule and the further configuration.

An Excel file can consist of several tables. The importer „reads“ every visible table. If the column heading of type „SS“ is not on the Table sheet is found, the sheet is not imported.

An import log gives information about the read files, tables and in which areas the tables were read and gives errors if necessary.

Figure 6 shows the example from the import log of the two source files from Figure 1 and 2 and the import rule.

7. In column A and B, the imported files and sheets are per line.

In column C and D is from which line and to which line the Import process has been performed. If a data check is activated in the import rule, a message indicating the cell address will be inserted in the output file if the check fails. In the example shown, this was in cell H12, see Figure 7.

The result of the import process is shown in Figure 8.

According to the optional configuration for indicating the origin of the data, the file name and the sheet name are included. The entry „* 600“ for „Facade Miller“ was not imported because this cell content does not correspond to the specified format. An alternative variant of the import rule generates a row-oriented output for the keywords with the type „S“, see Figure 9.

For such an output, the type code „SZ“ is assigned in the import rule for the keywords. The changes in the import rule are highlighted in yellow in Figure 11.

The data import writes the results into a normal Excel sheet, but the number of rows is limited. If the import process exceeds this limit, multiple spreadsheets will be created.

The dialog interface for the displayed import function is freely configurable. The individual ideas abstractly refer to the intended function.

With the introduced import functionality data for one or more files can be read controlled from tables.

The systematic and automatic process avoids a manual one. Compilation error. Unclean data can be checked during import. The import process is logged.

Compared to a manual compilation creates a great time advantage.

With the optional version of the line-oriented output, the data can be processed more easily.

This also applies to Excel as well as to third-party systems. For example, the import result can be saved on the Excel side as a „csv“ file format (considered as a universal file format).

The technical implementation of this import functionality is exemplified using the tool „Excel Collector“.

Process example:

Installation of the import regulation

The Excel tool is currently designed so that the import rules are created in the tool's own worksheets. If another import rule is to be created, copy the worksheet. The import process is started with the program button „Import Excel files according to import rule“.

On the left in Figure 11 is a simple description; on the right in Figure 11 is the import rule.

After the call „Import Excel files according to import rule“ a selection dialog for the files to be read appears, see Figure 12. Individual files or folders can be selected, see Figure 13.

As a result, all Excel files in the folder are transferred to the import list, see Figure 14. In the selected example the file „Tab_3“ should be removed from the import list. This is implemented using the corresponding „Remove selected entries“ button.

Possibly. Further options of the import are activated. In the example in Figure 15, the output of the file name is selected and then the import is started.

presentation of results

After the import, a message will appear as shown in Figure 16.

The import result is in a new Excel file under the „Import“ worksheet, see Figure 17.

The import log is stored as a separate table (eg „Script 04.01.18“), see Figure 18.

Advantages:

- Possibility of controlled reading of data from one or more tables.
- Avoiding incorrect entries compared to manual copying.
- Possibility of checking dirty data.
- Report creation of the process.
- Significant time and effort savings compared to manual copying.
- Easier processing of the data in variant with line-oriented output.
- Ability to save the import result on the Excel side as „csv“ file format.

Possible application:

- Applicable in spreadsheet programs.

Stand der Technik

	A	B	C	D	E	F	G	H
1								
2								
3								
4		Projekt	Projekt-ID	2016	2017	2018	2019	
5								
6		Aktuelle Projekte						
7		Polsterung "Hauser"	5910		100	500		
8								
9		Streichen "Herter"	5911			600	100	
10								
11		Schreinern "Huber"	5912			200		
12								
13		Zusatzinformation						
14								
15		Alle Projekte von H. Müller						

Abbildung 1

	A	B	C	D	E	F	G	H	I
1		Projekte H. Heinze							
2		Projekt	Projekt-ID	RechNr	erhalten	2016	2017	2018	2019
3									
4		Alte Projekte							
5		Zaun "Heinrich"	4711	4711-0	ja	2000	500		
6		Fassade "Müller"	4712	4712-0	nein		5000		
7									
8		Aktuelle Projekte							
9		Gartenteich "Reiter"	4910			300	2000		
10		Pflaster "Zeilinger"	4911				8000	2000	
11									
12		Fassade "Müller"	4912				*600		

Abbildung 2

	A	B	C	D	E
1	Projektnummer	Projektname	2017	2018	2019
2	5910	Polsterung "Hauser"	100	500	
3	5911	Streichen "Herter"		600	100
4	5912	Schreinern "Huber"		200	
5	4910	Gartenteich "Reiter"	300	2000	
6	4911	Pflaster "Zeilinger"		8000	2000
7	4912	Fassade "Müller"		*600	

Abbildung 3

Technische Neuerung

	B	C	D	E	F	G	H
Import - Vorschrift							
	Schlüsselwort [Text,Zahl]	Typ [SS, S,SZ, KB,KE]	Spaltenüberschrift Output	Zeilenbenennung bei Typ SZ	Spaltennummer im Output [1..]	Datentyp [Datum, Zahl, leer]	Filter bei SS Erläuterung
Projekt		SS	Projektname		2		
Projekt-ID		S	Projektnummer		1		
2017		S	2017		3	Zahl	
2018		S	2018		4	Zahl	
2019		S	2019		5	Zahl	
2020		S	2020		6	Zahl	
Aktuelle Projekte		KB					
Zusatzinformation		KE					

Abbildung 4

	A	B	C	D	E	F	G	H
1								
2								
3								
4		Projekt	Projekt-ID	2016	2017	2018	2019	
5								
6		Aktuelle Projekte						
7		Polsterung "Hauser"	5910		100	500		
8								
9		Streichen "Herter"	5911			600	100	
10								
11		Schreinern "Huber"	5912			200		
12								
13		Zusatzinformation						
14								
15		Alle Projekte von H. Müller						

Abbildung 5

	A	B	C	D	E	F
1	Importierte Datei	Tabellenblatt	Zeile von	Zeile bis	Zeilen	Fehler
2	M:\Doßke\Patent\Sammler\Tab_1.xlsx	Tabelle2	7	12	3	
3	M:\Doßke\Patent\Sammler\Tab_2.xlsx	Tabelle1				Datentyp stimmt nicht mit Zielformat überein in SHS12
4			9	12	3	
5	M:\Doßke\Patent\Sammler\Tab_2.xlsx	Tabelle3				0 Tabellenblatt leer
6	M:\Doßke\Patent\Sammler\Tab_2.xlsx	Tabelle4				0 Tabellenblatt leer

Abbildung 6

	A	B	C	D	E	F	G	H	I
1		Projekte H. Heinze							
2		Projekt	Projekt-ID	RechNr	erhalten	2016	2017	2018	2019
3									
4		Alte Projekte							
5		Zaun "Heinrich"	4711	4711-0	ja	2000	500		
6		Fassade "Müller"	4712	4712-0	nein		5000		
7									
8		Aktuelle Projekte							
9		Gartenteich "Reiter"	4910			300	2000		
10		Pflaster "Zeilinger"	4911				8000	2000	
11									
12		Fassade "Müller"	4912				*600		

Abbildung 7

Technische Neuerung

	A	B	C	D	E	F	G	H
1	Projektnummer	Projektname	2017	2018	2019	2020	Dateiname	Blattname
2	5910	Polsterung "Hauser"	100	500			Tab_1.xlsx	Tabelle2
3	5911	Streichen "Herter"		600	100		Tab_1.xlsx	Tabelle2
4	5912	Schreinern "Huber"		200			Tab_1.xlsx	Tabelle2
5	4910	Gartenteich "Reiter"	300	2000			Tab_2.xlsx	Tabelle1
6	4911	Pflaster "Zeilinger"		8000	2000		Tab_2.xlsx	Tabelle1
7	4912	Fassade "Müller"					Tab_2.xlsx	Tabelle1

Abbildung 8

	A	B	C	D
1	Projektnummer	Projektname	Jahr	Wert
2	5910	Polsterung "Hauser"	2017	100
3	5910	Polsterung "Hauser"	2018	500
4	5910	Polsterung "Hauser"	2019	
5	5911	Streichen "Herter"	2017	
6	5911	Streichen "Herter"	2018	600
7	5911	Streichen "Herter"	2019	100
8	5912	Schreinern "Huber"	2017	
9	5912	Schreinern "Huber"	2018	200
10	5912	Schreinern "Huber"	2019	
11	4910	Gartenteich "Reiter"	2017	300
12	4910	Gartenteich "Reiter"	2018	2000
13	4910	Gartenteich "Reiter"	2019	
14	4911	Pflaster "Zeilinger"	2017	
15	4911	Pflaster "Zeilinger"	2018	8000
16	4911	Pflaster "Zeilinger"	2019	2000
17	4912	Fassade "Müller"	2017	
18	4912	Fassade "Müller"	2018	
19	4912	Fassade "Müller"	2019	

Abbildung 9

Import - Vorschrift

Schlüsselwort [Text,Zahl]	Typ [SS, S,SZ, KB,KE]	Spaltenüberschrift Output	Zeilen- benennung bei Typ SZ	Spalten- nummer im Output [1..]	Datentyp [Datum, Zahl,leer] Filter bei SS	Erläuterung
Projekt	SS	Projektname		2		
Projekt-ID	S	Projektnummer		1		
2017	SZ	Jahr	2017	3	Zahl	
2018	SZ	Jahr	2018	3	Zahl	
2019	SZ	Jahr	2019	3	Zahl	
2020	SZ	Jahr	2020	3	Zahl	
Aktuelle Projekte	KB					
Zusatzinformation	KE					

Abbildung 10

Technische Neuerung

Beispielablauf:

ExcelSammler: automatisches einlesen von Excel-Dateien
Wolfgang Doßke

Exceldateien nach Import - Vorschrift einlesen

Funktion:
Im wählbaren Ordner werden alle Tabellenblätter aller Excel-Dateien ausgelesen und in ein Tabellenblatt geschrieben. Beim Importvorgang wird ein Skript-Protokoll erzeugt. Der Importvorgang wird über Schlüsselwörter im Vorschriftenblock gesteuert (s. rechts).

Der **Überschriftsbereich** mit den Typen SS, S und SZ dürfen maximal zwei Leerzeilen auseinanderliegen. Es muß mindestens ein Schlüsselwort vom Typ SS existieren. Importiert wird in der Spalte, in der das SS-Schlüsselwort steht.

Der **einzelnde Bereich** kann über die Schlüsselwörter vom Typ KB und KE **eingeschränkt** werden. Fehlt der Kenner KB so wird direkt nach dem Überschriftblock importiert. Fehlt der Kenner KE, so wird bis zum Ende importiert.

Darüber hinaus ist optional ein **variabler Importfilter** für das **Schlüsselwort** vom Typ SS anlegbar (s. Datentyp). Gültig sind die typischen Regeln: ? für beliebiges Zeichen, * für beliebige Zeichenfolge, # für Ziffer, ...

Sofern auf ein gewünschtes Format geprüft werden soll, ist dieser Datentyp mit anzugeben (Leer bedeutet beliebiges Datenformat). Schlägt die Prüfung fehl, wird im Skript-Protokoll ein **Fehlerbericht** eingetragen.

Die erzeugte Excel-Datei kann im csv-Format manuell abgespeichert werden.

Schlüsselwort [Text,Zahl]	Typ [SS, S,SZ, KB,KE]	Spaltenüberschrift Output	Zeilen- benennung bei Typ SZ	Spalten- nummer im Output [1..]	Datentyp [Datum, Zahl,leer] Filter bei SS	Erläuterung
Projekt	SS	Projektname		2		
Projekt-ID	S	Projektnummer		1		
2017	S	2017		3	Zahl	
2018	S	2018		4	Zahl	
2019	S	2019		5	Zahl	
2020	S	2020		6	Zahl	
Aktuelle Projekte	KB					
Zusatzinformation	KE					

Abbildung 11

Import von Excel-Dateien nach Importvorschrift

Ordner auswählen Liste leeren

Einzelne Datei auswählen markierte Einträge entfernen

Einstellungen Import: Einstellungen Ausgabe:

Formel "Telergbnis" nicht importieren Ausgabe Dateiname

Formel "Summe" nicht importieren Ausgabe Blattname

Import starten

Abbildung 12

Technische Neuerung

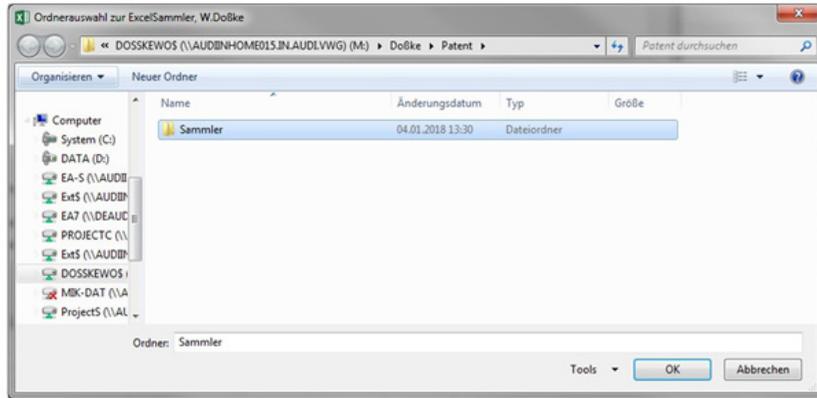


Abbildung 13

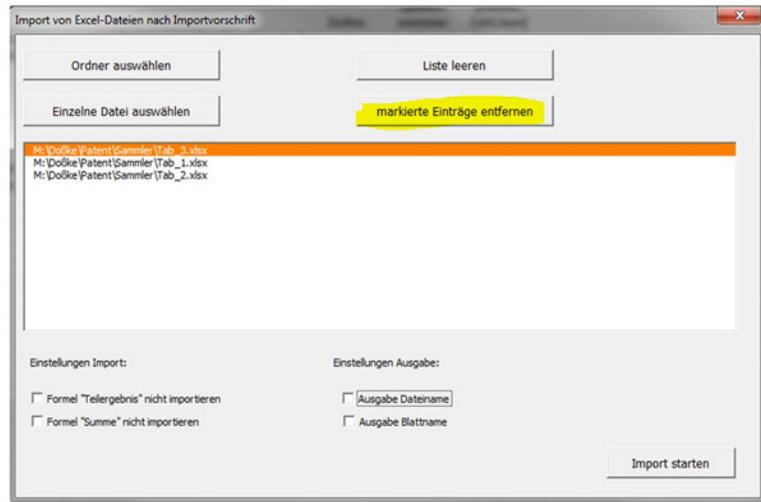


Abbildung 14

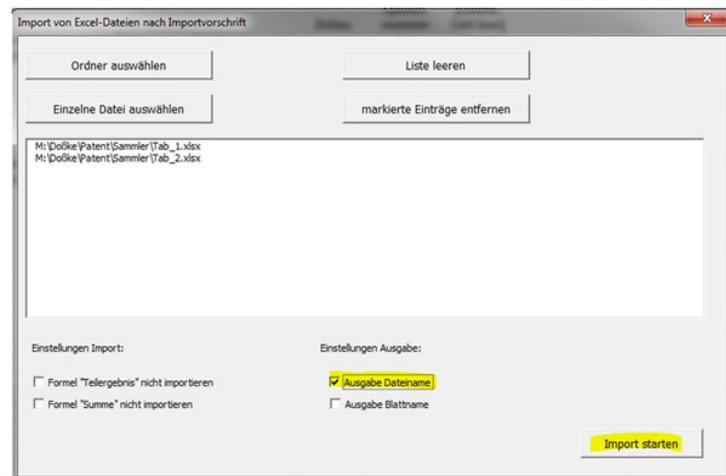


Abbildung 15

Technische Neuerung

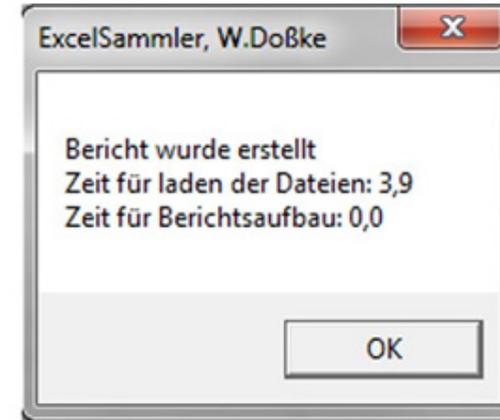


Abbildung 16

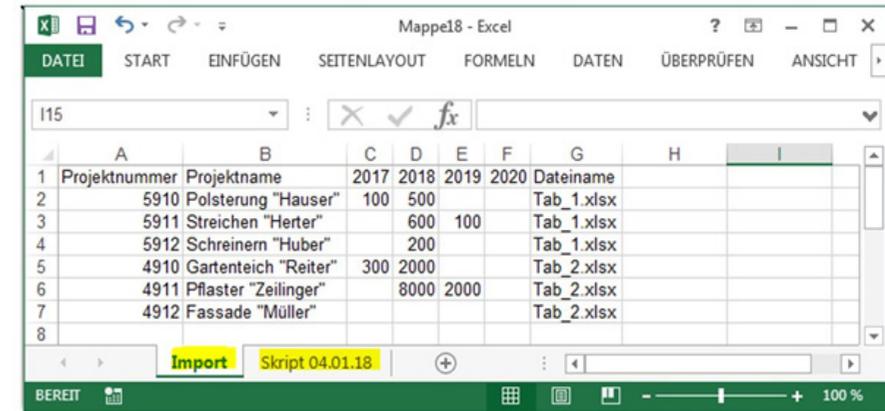


Abbildung 17



Abbildung 18