Archiving Data

Food for thought and technical possibilities

FAQ
9 December 2021
Created with Version 13.0.5.1

Information about this document

Parts of this document may be automatically translated.

Document History

Version	Date	Author(s)	Modifications / Remarks
	06.10.2021	GA	Initial Release
	13.10.2021	GA	Translation

CONTENTS

1 Foreword

This document is intended to provide some food for thought on the subject of archiving. Unfortunately, the topic of archiving only reaches most customers when the case has already arisen that their database has too many files for daily operation. However, the topic should be planned from the beginning when starting a project.

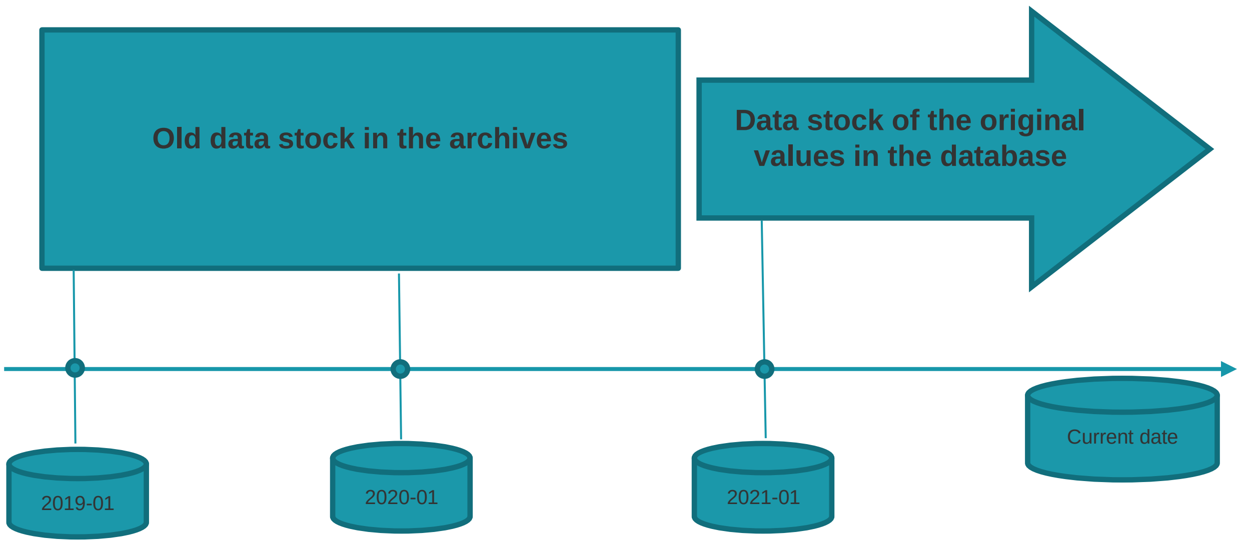
The project team also offers on-site support in the form of workshops. You can find the link to this on the homepage.

Let's look at a timeline. Every company has the obligation to keep data. Two to 50 years, depending on the area in which a company operates and the legal requirements. However, this data does not have to be in one big database.

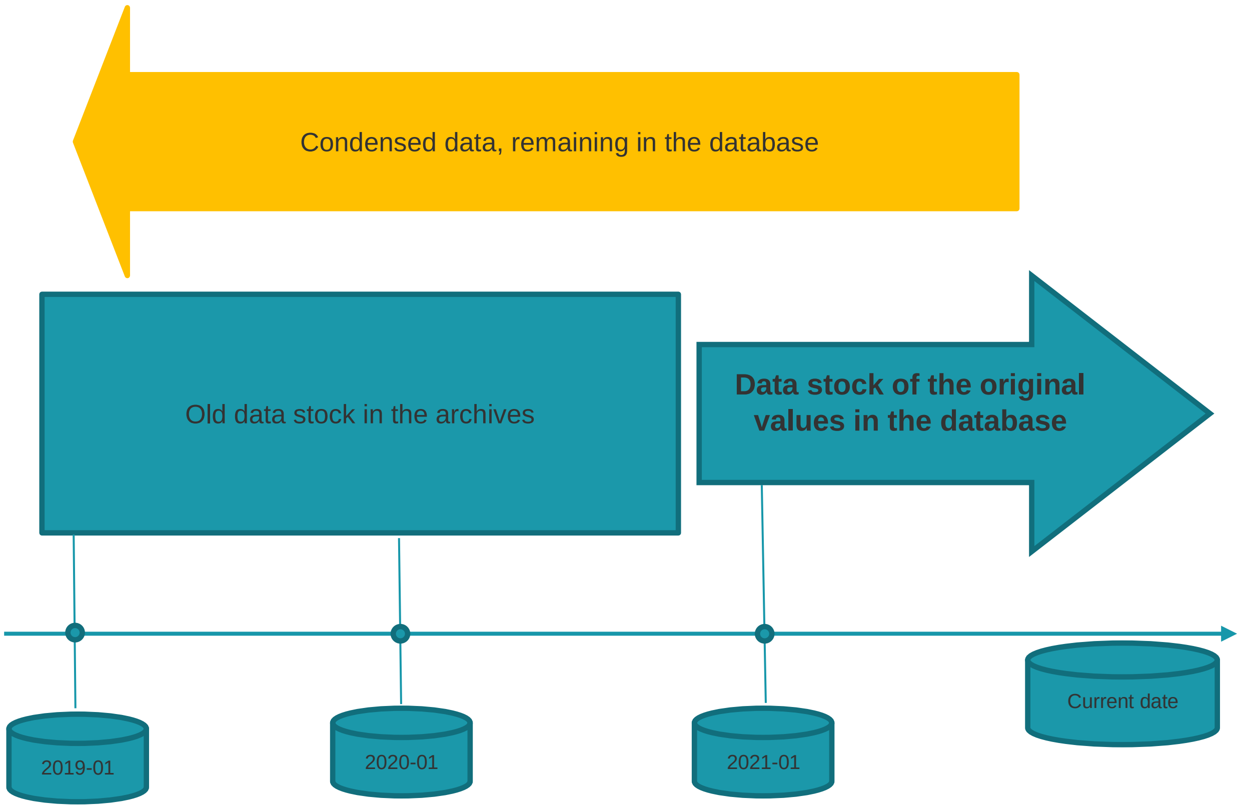
The consideration is: which data stock must always be directly available in the database for current evaluations, for daily business. But at what age can users be expected to access data stored in archives?

Let's look at the following example:

As an example for consideration: with the current date **2021-10,** the company decides to only have data from the last 14 months in the live DB. So everything older than 14 months can be archived



With the M-QIS product, data can also be compacted BEFORE archiving. The compaction saves the calculated characteristic values, which can then be read without measured values in the long-term analysis module. Thus, in addition to the benchmarks and the development of the processes over time, a condensed view of older data would also be possible via this module.



2 Preliminary considerations for archiving

At the beginning of any project with the Q-DAS software, or at the beginning of archiving, the following points must be considered.

		Comment
Time range	The time domain consideration includes 2 points: • Data back to what time range should be kept in the main database for direct analysis? • Into which sub-areas (month / quarter / year) should the archive packages be divided?	The data stock in the main database should correspond to what needs to be continuously accessible in day-to-day business. The size of the archive units depends largely on the archive type and the amount of data per unit. Each archive type has a natural limitation in terms of size.
Archive data type	The following are available • dfq files • Access databases • SQL/Oracle databases	Each type of archive has its advantages and disadvantages. These are listed in chapter 3.
Manual / automatic	Manual archiving is available in every Windows application. Automatic archiving is available in the M-QIS product	Manual archiving is to be regarded as an emergency solution; a recommendation, especially in the case of a constant flow of data, is the automatic removal of obsolete original values.
Data compression before outsourcing	With the M-QIS product, data compression can take place before outsourcing.	The data compression for the Long Time Analysis module offers the possibility of still being able to view the calculations even if no more original values of this time range are available. This is only useful when using MS SQL or Oracle databases.

3 Archive types

Depending on the archiving tool, more or fewer archive types are available. Their occurrence as well as their advantages and disadvantages will be explained here.

3.1 DFQ

Archiving to DFQ files is available in all versions and is the preferred automated archiving option. DFQ files are the main format of the Q-DAS software and will always be readable even with later versions.

One thing to be concerned about when archiving in DFQ files is the information that will be lost, as this is only kept in databases:

- Part measurement status
- Assigned documents
- PDCA information
- Catalogue information
- Additional data-specific quality control charts
- Change history
- Mother / Child - Structures

In most cases, however, there is no obligation to archive this information, so it can be disregarded.

3.2 Access

Archiving in Access databases is a historical piece of software. It no longer has any significance in new projects.

When archiving in Access databases, the Access databases can then be accessed with the software in read-only mode. The problem is the database size, Access is limited to 2GB, so the amount for an archive here must be considered in advance and this must also be tested.

Archiving in Access databases is only possible with manual archiving.

3.3 SQL/Oracle

This is a special case. The archive can be archived in the same MS SQL/Oracle database each time. All tables are extended with the current archive number.

These archives cannot be accessed via the software. If the data is needed, it is only possible to restore the archives to the main database.

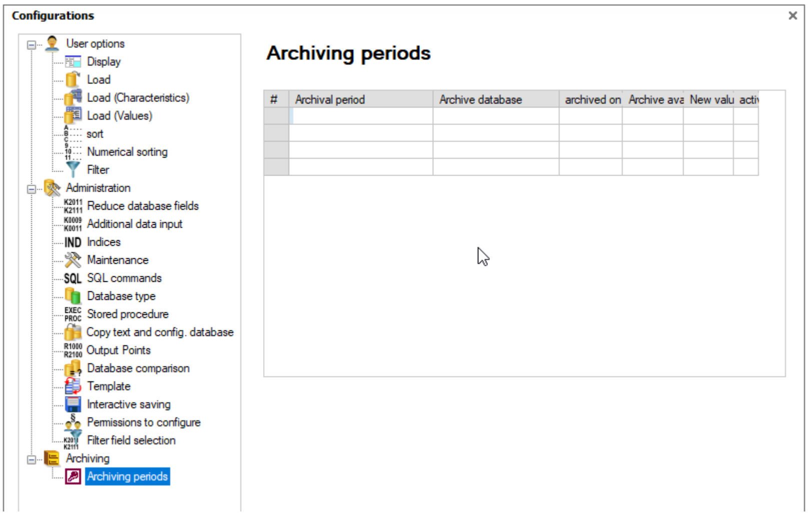
Archiving in MS SQL / Oracle databases is only possible with manual archiving.

4 Technical types of archiving

2 different archiving tools are available. Manual archiving (this is possible in every standard application), and automatic archiving as part of M-QIS.

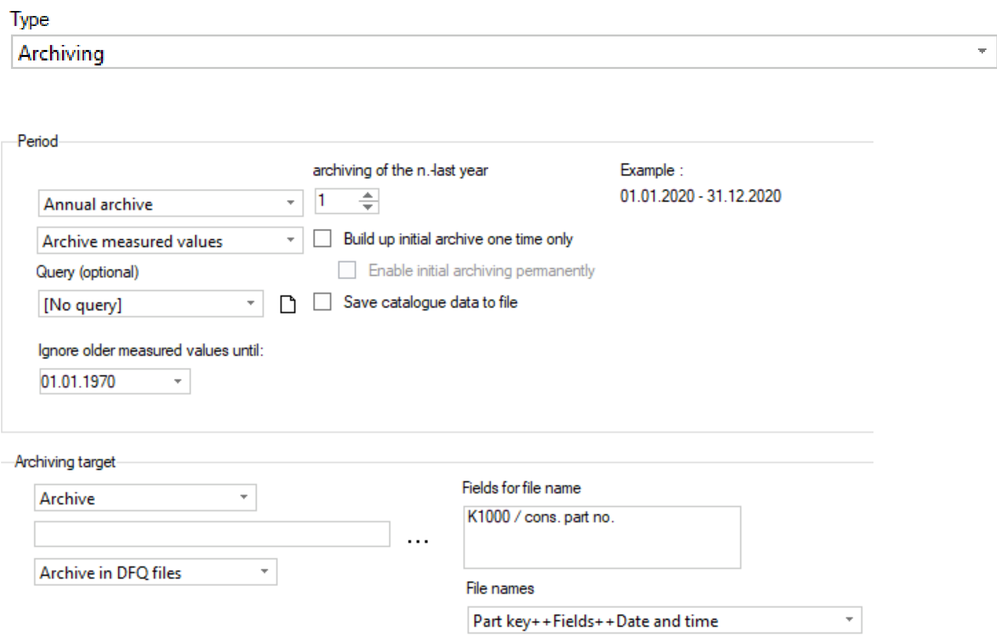
4.1 Manual archiving

Manual archiving can be called up in the database options, provided the database administration authorisation is given. :



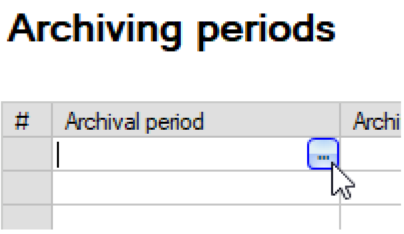
4.2 Automatic archiving in M-QIS

In M-QIS Engine Reporting, this is available with the type "Archiving

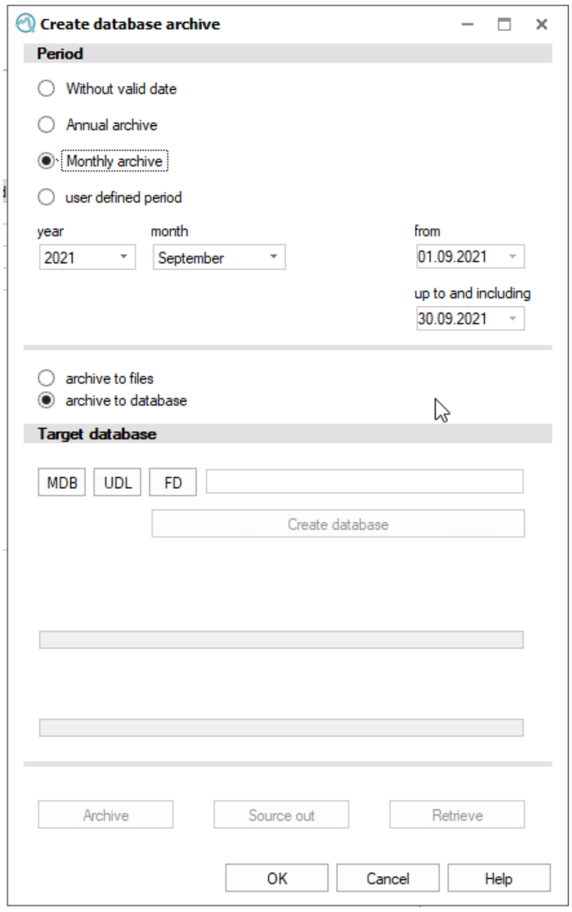


5 Manual archiving procedure

In the archive option, click in the first free cell, then the menu can be opened:



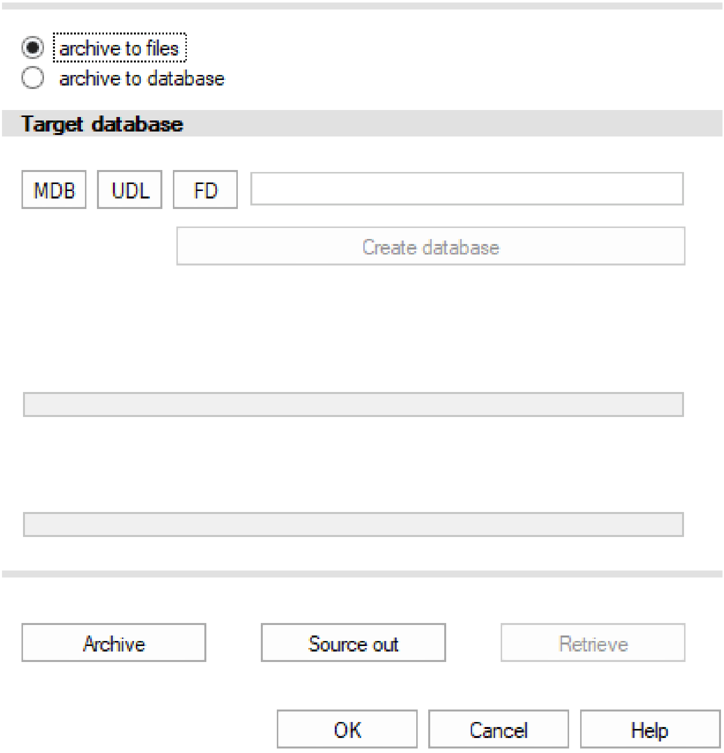
The following dialogue then opens:



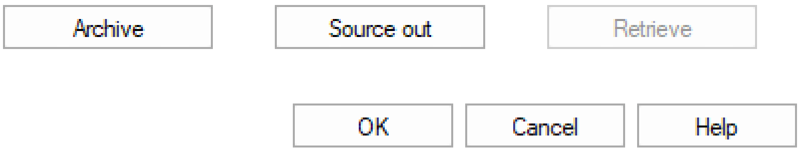
In the upper part, the range of the period to be archived is specified. Depending on the target type, the size here must be chosen to be correspondingly small (mainly if the target is an Access archive database).

5.1 Archive to files" target type

With the target type "archive to files", the database connection below no longer has any effect.



5.1.1 Archive or outsource

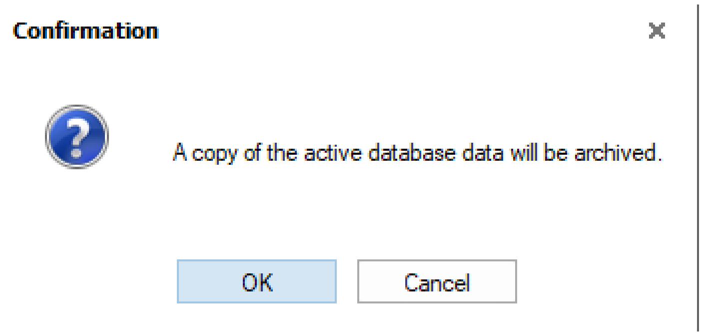


The archive is started with the two options "Archive" or "Store".

The difference between the two options is:

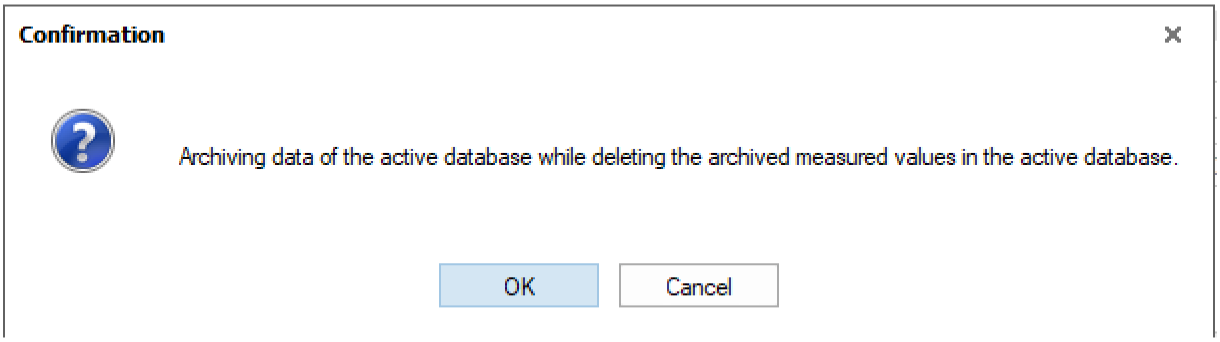
Archive

You will be informed that this will only create a COPY. No data is outsourced



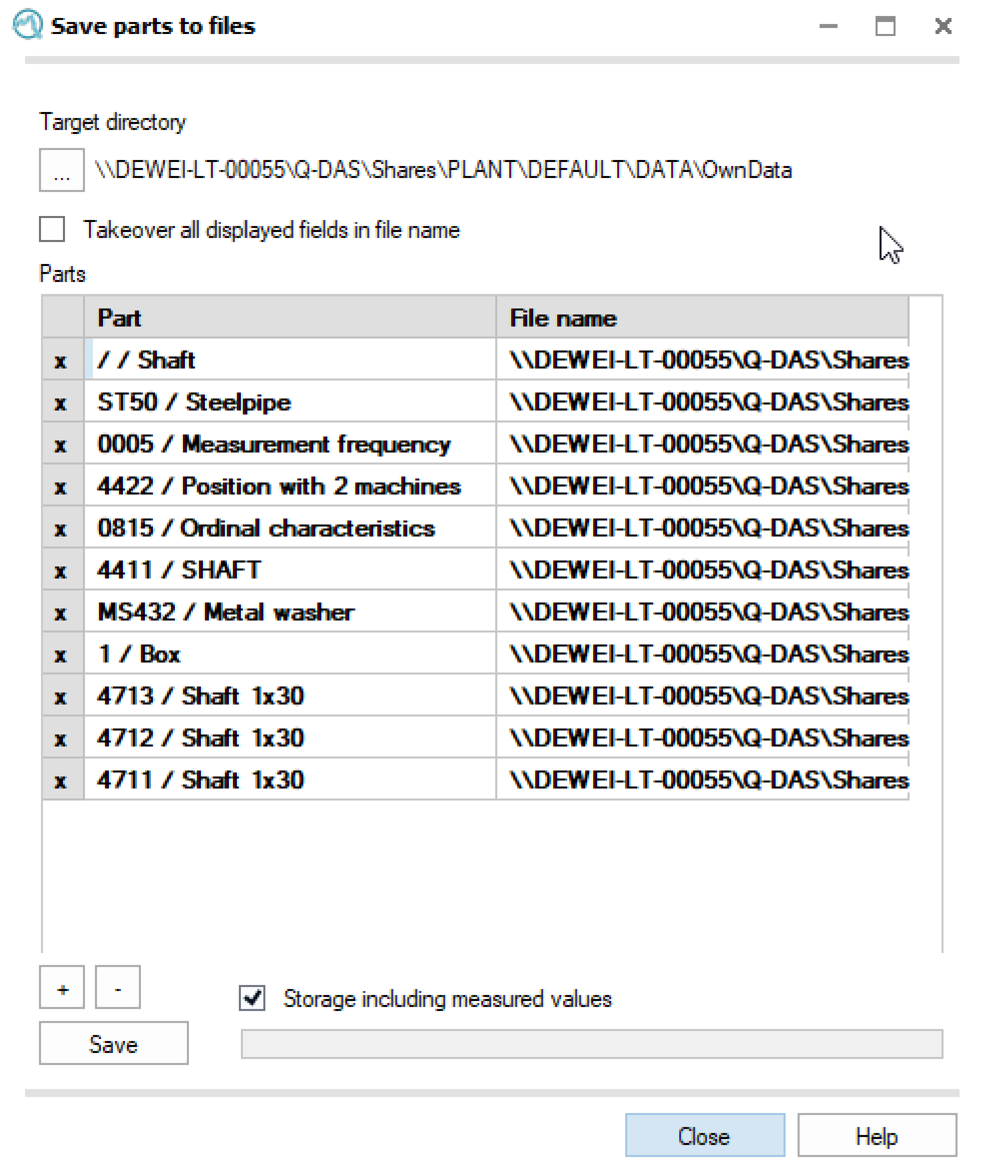
Outsource

A message appears informing you that the measured values that are then in the archive will now also be removed from the productive database.



In the archive, a dfq file is created for each part of the database. Whether the dfq contains values is then decided only by the time range.

After pressing one of the two options, the dialogue "Save parts to file" opens.



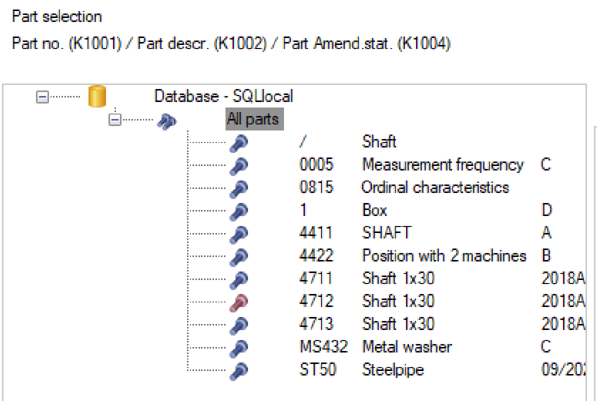
All parts are pre-marked. If desired, individual parts can be deactivated here by clicking on the "x".

Target directory

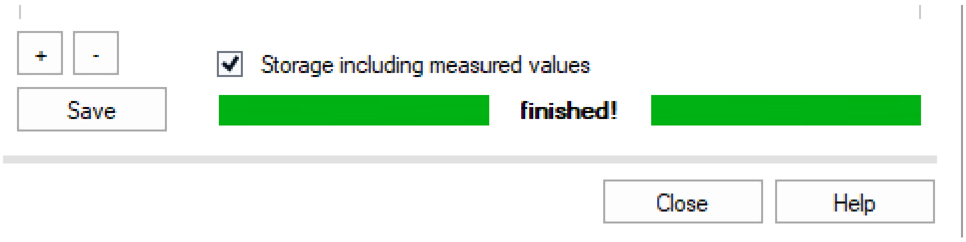
Initially, the standard directory for storing own files is specified. Optionally, a separate directory can be selected.

Take over all displayed fields in file names

By default, only K1001 and K1002 are assigned in the file names in addition to the internal (and thus unique) consecutive part number. With the said option, all fields that are specified in the dialogue "Read from database" are included in the file name.



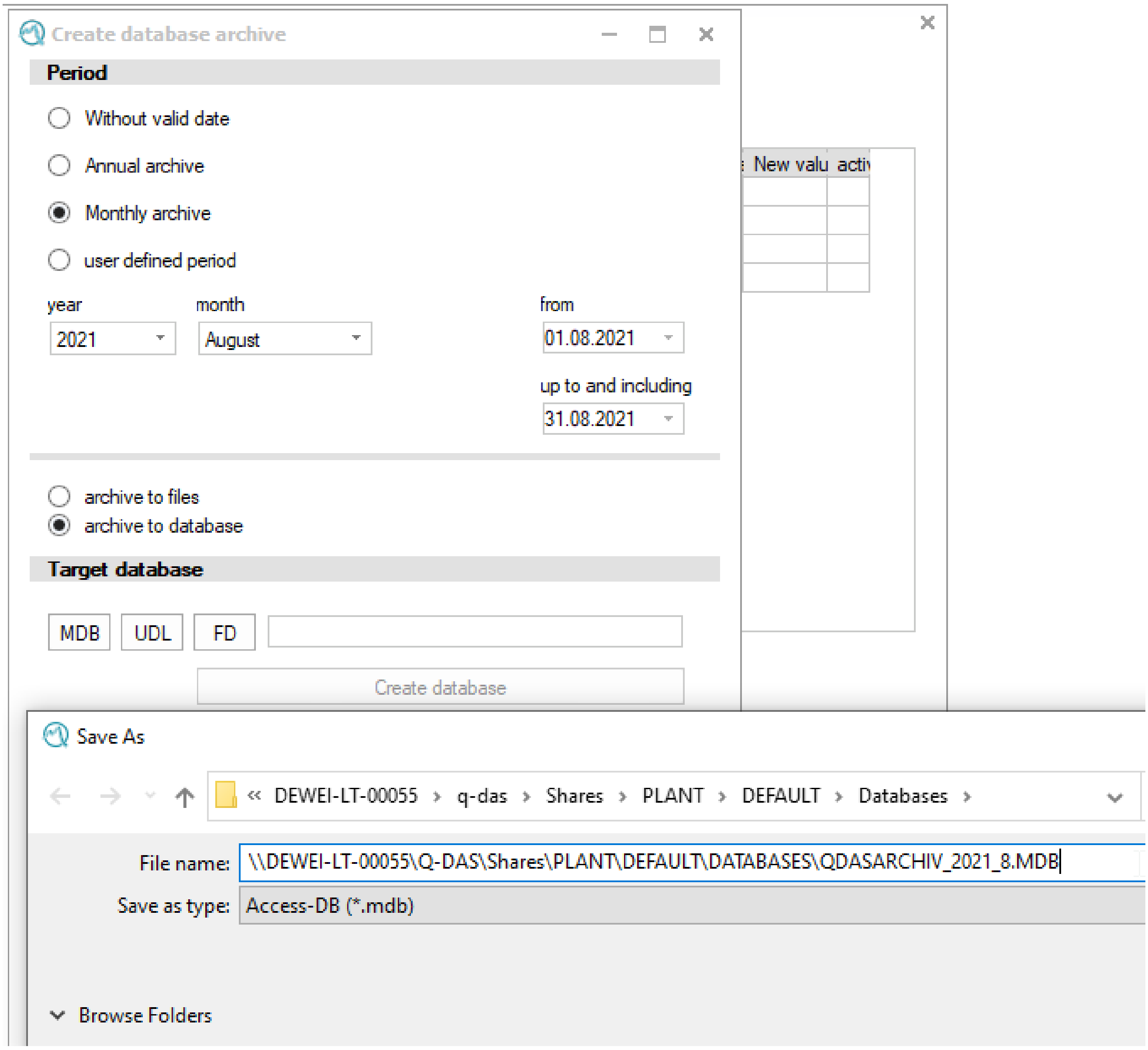
Click on "Save" to start the archiving process. After completion, this is displayed:



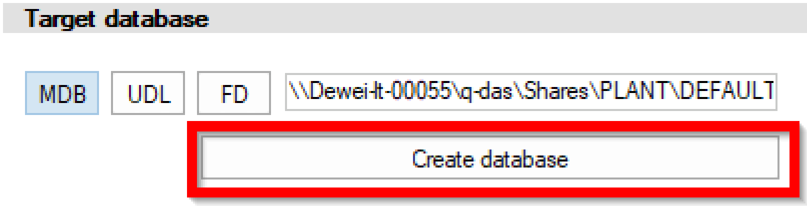
5.2 Target type "Archive to database MDB”

If "Archive to database" is selected as the target type, the database type must then be selected.

If "MDB" is selected, the file name is suggested according to the date specifications.



The database must then be created with "Create database":



5.2.1 Archive or outsource

As with DFQ files, there is a difference between archiving (making a copy) and swapping out (removing from the main database).

In the archive database, with the exception of the WERTEVAR table (which contains the measured values), all other tables are unique copies of the main database.

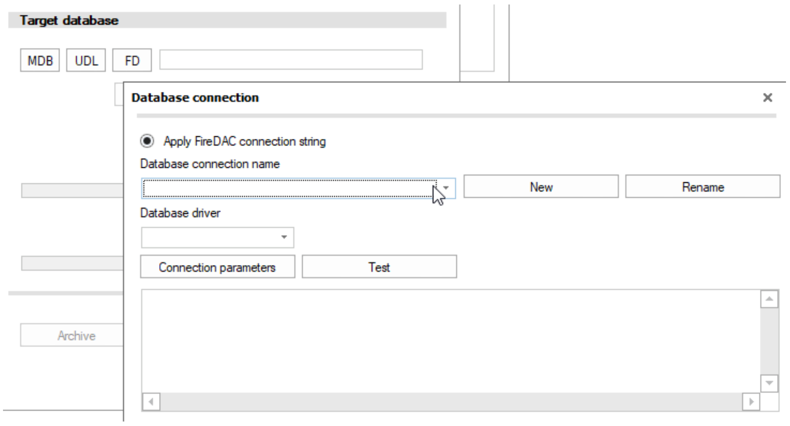
The archive MDB can be accessed directly, but only read. The Access DB is write-protected.

5.3 Target type “Archive to database FD"

If "Archive to database" is selected as the target type, the database type must then be selected.

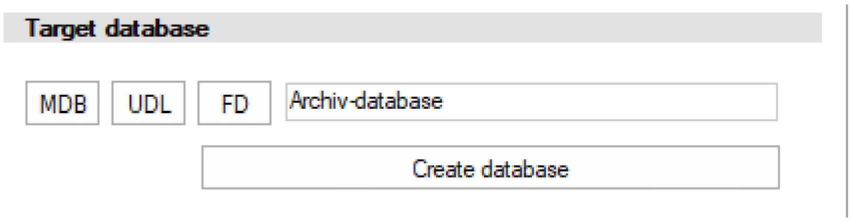
If "FD" is selected, a firedac connection to the SQL / Oracle database is created:

None of the classic database connections are shown in the dialogue. A new connection must be created here:

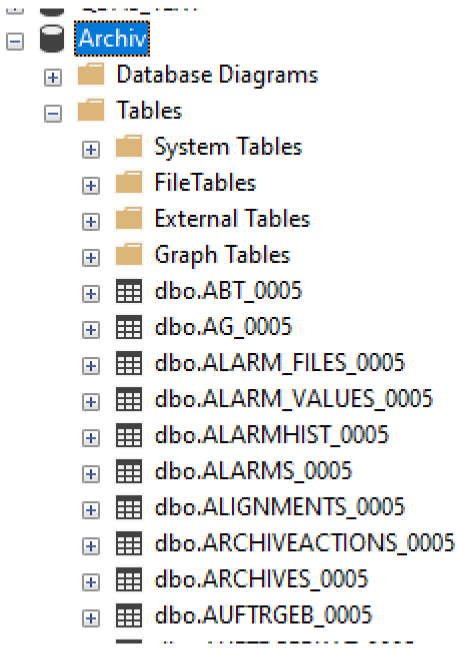


The first database must be a new, empty archive database. (not an existing one!)

The tables for the archive are then prepared with "Create Database".



The table of an archive database (MS SQL / Oracle) has the sequential number of the archive as a prefix. This means that any archives can be stored in an archive database.



However, this also means that direct access to an archive database is not possible.

5.3.1 Archive or outsource

As with DFQ files, there is a difference between archiving (making a copy) and swapping out (removing from the main database).

In the archive database, with the exception of the WERTEVAR table (which contains the measured values), all other tables are unique copies of the main database.

The archive database (MS SQL / Oracle) can NOT be accessed directly

6 Archiving procedure in the reporting system

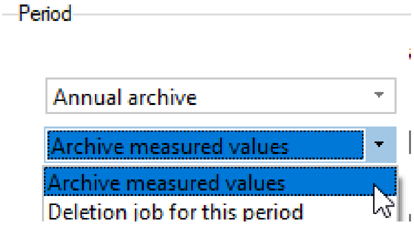
In the reporting system, this is created as an "archiving" job:

It is not possible to describe all the options in their combination, so the various use cases are discussed here

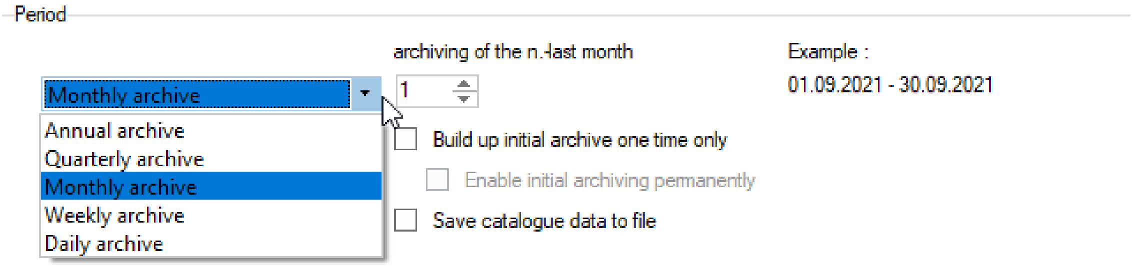
Archiving by the M-QIS Reporting System is only possible in *.dfq files.

6.1 Use Case: Archiving / outsourcing values

To archive / outsource measured values, select "Archive measured values":

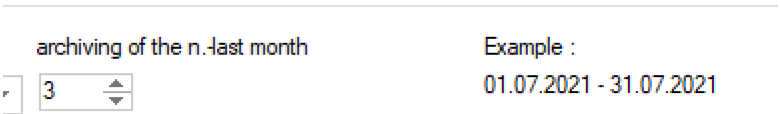


Now the time ranges are defined. You can choose from daily to yearly ranges.



6.1.1 Archiving the n. - last time unit

This is the indication of which dates are still to be kept. For example, the current day is 13 October 2021. The following is selected: the 3rd to last month is to be archived; this is also shown in the example to the right.

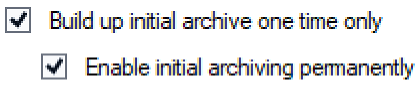


6.1.2 Initial archiving

If archiving is started with the M-QIS Reporting System, there is usually a much older dataset than the one actually specified in the time range. With the option "Perform one-time initial archiving", all older time ranges are also archived in the same unit.

The option is volatile and would be deactivated again after the first run of the job.

If the option is to be set permanently (e.g. because there is a suspicion that old data keep coming into the database through feeding systems), the second option is also set, then this option remains.



6.1.3 Save catalogues to file

If there is a suspicion that the catalogues change permanently over time, this option can be used to save each individual DFQ file with the catalogue to file. (K4xxx fields)

When opening such a file, the software switches directly from the current catalogue setting to "Catalogues from dataset".

6.1.4 Ignore older measured values until

An end date can be specified here, up to which back initial archiving should check the date.

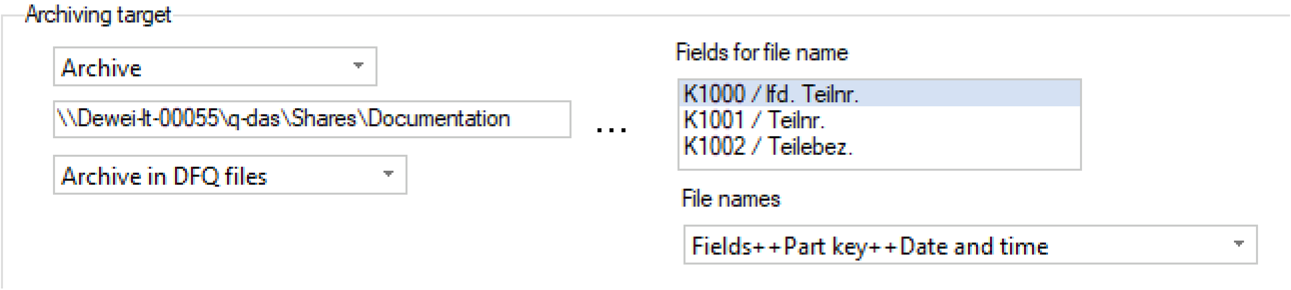
The reason for the existence of this option is as follows: If measurement-generating systems had not written the date and time for values, or had written them incorrectly, and the upload was not set to correct this, then values from 01.01.1970, the 0-time, can exist. In order to avoid monthly archives without content from 1970 to the current year, a maximum date can be set here.

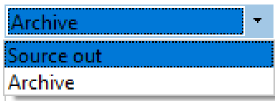
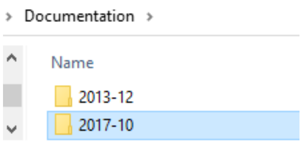
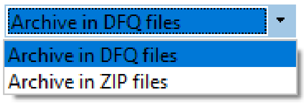
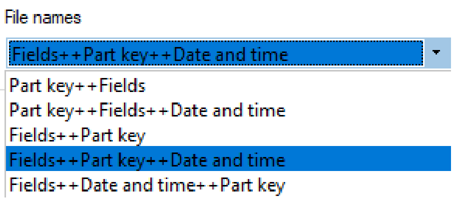
In a cleanly set up system, this option should be of no significance.

6.1.5 Archiving target

This section specifies in which form and where the files are placed.

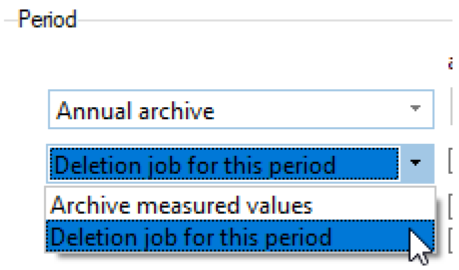
In addition to the basic settings, as with manual archiving, further specifications are possible here, which are explained in the following table.



Archive / Source out  	As with manual archiving, a decision can be made here between archive (copy) and effective retrieval.
Destination	A main folder must be selected here. Subfolders are created in this folder depending on the selected time unit. Example monthly archives 
File type  	Due to the high compression rate of dfq files, it is recommended to store zipped dfq files as the file type. With the software, a zipped dfq file can also be opened directly with the option "Open from zip".
Field names	Fields for file name generation can be selected here. 
File name generation	The file name itself always contains the partkey, the internal key, as only this is unique in the database. With the selection here, the order of the names can be determined. Date/time are the current time, i.e. when the file was created. 

6.2 Use Case: Delete data

To delete "data", "Delete job for this period" is selected:



This is an **irreversible action!** Only in consultation with the local IT and cyclical backups of the database should this option be used!

"Data" is generally:

- Measured values
- Saved results
- Alarms
- Signatures

The setting of the time range is analogous to archiving / swapping.

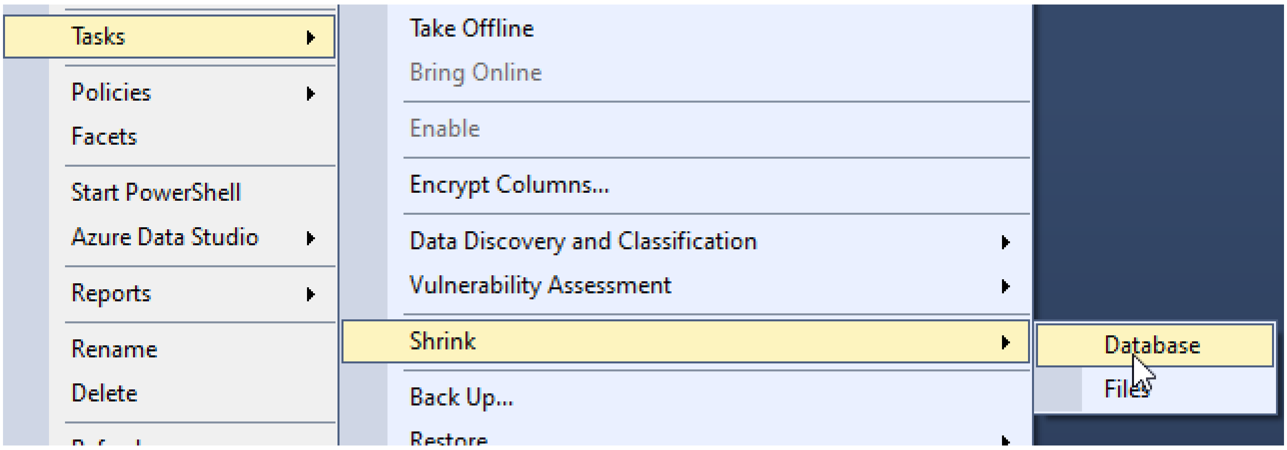
Delete measured values	Analogue to the swapping, the measured values are only deleted here without saving them beforehand.
Delete results	Saved results are the compressed data of the long-term analysis, which were summarised with the reporting system (division allocation for saving the results). It is also possible to select only data of a certain division allocation.
Delete alarms	Alarms are stored in the database either via the upload (single value alarms) or via the procella.
Delete signatures	Signatures are stored in the database, e.g. via the report view or via automatisms in procella.

7 Annexes

When data has been removed/deleted for the first time due to overflowing databases, it is advisable to think about various other "clean-ups".

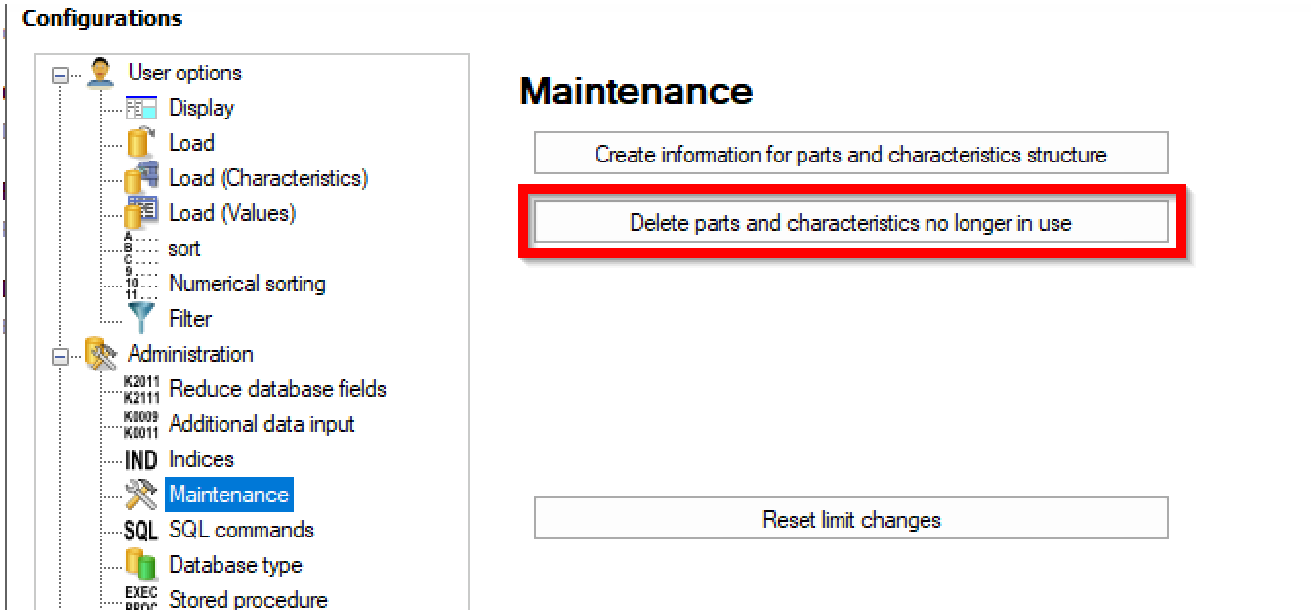
7.1 Shrink or compress and repair database

After major clean-ups, the databases should be "compressed". In the case of MS SQL databases, for example, this would be the task "Skrink".

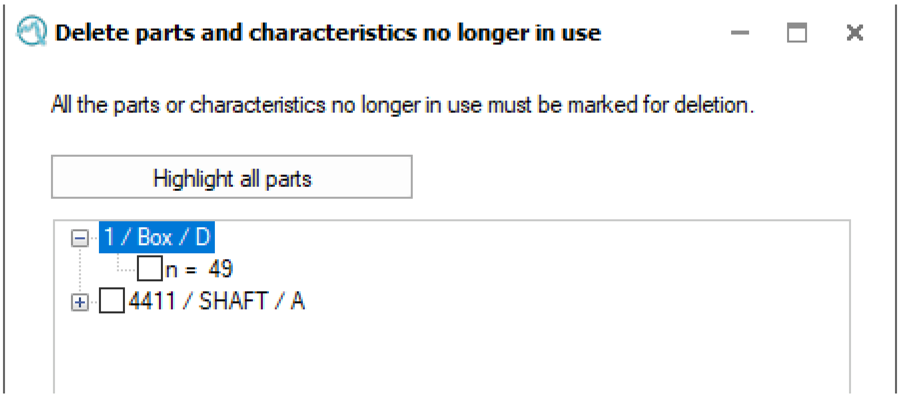


7.2 Deleting parts/features that are no longer used

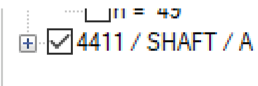
This question comes up automatically after manual archiving, but can be started at any time via the database options



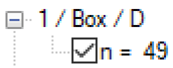
After pressing the button and searching, the following dialogue is displayed:



All parts that no longer have any measured values are displayed in all characteristics, to be marked directly BEFORE the part:



If only certain characteristics are empty for a part, the characteristics could still be marked individually, but not the entire part:



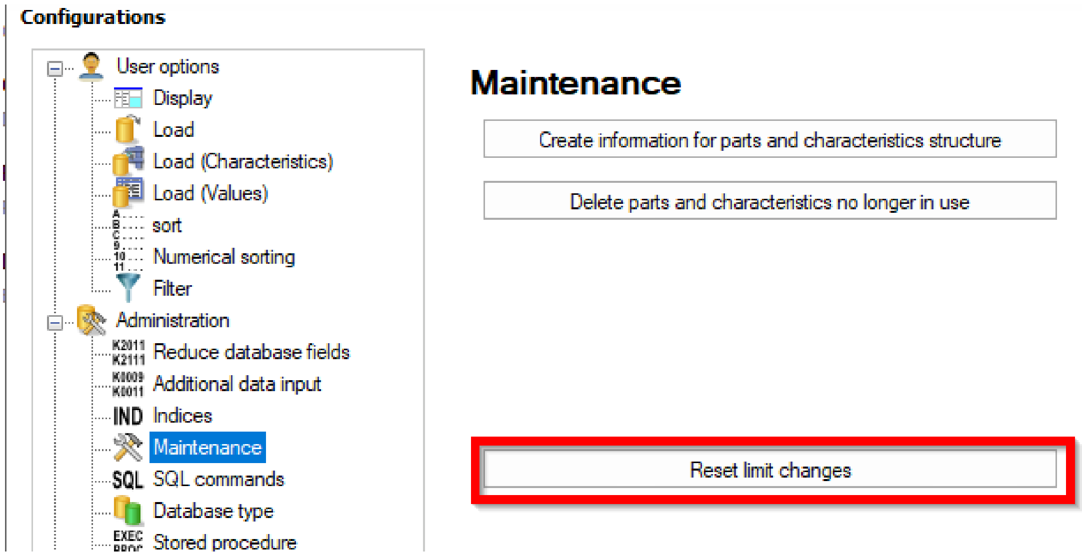
The option "Mark all parts" would only mark completely empty parts.

If the database was archived in databases, there may not be any empty parts! If the last measured value of a part was archived, a "placeholder measured value" is written that cannot be seen via the software itself. This serves to ensure that the option "find empty parts" does NOT find these parts, as the internal consecutive part number must be retained for possible restorations.

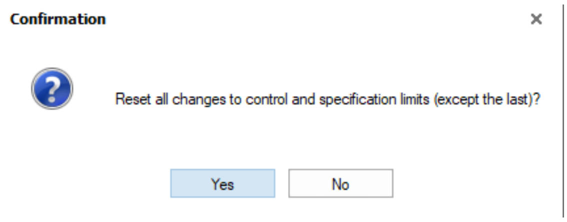
This should only be done after a backup of the database by the local IT!

7.3 Reset change of limits

On the same tab is the option "Reset change of limits".



Brief explanation: Even without the change history activated, every change to specification limits or action limits is saved. With this option, the for all records in the database can be cleaned up so that only the last status (which the software would also read directly) exists.



A detailed selection is not possible here!

This should only be done after a backup of the database by the local IT!