Archivierung von Daten

Ein Denkanstoß und technische Möglichkeiten

FAQ
9 December 2021
Created with Version 13.0.5.1

Information about this document

All rights, including translation in foreign languages, are reserved. It is not allowed to reproduce any part of this document in any way without written permission of Hexagon.

Parts of this document may be automatically translated.

Document History

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

CONTENTS

1 Vorwort 4

2 Vorüberlegungen zur Archivierung 6

3 Archiv-Typen 7

3.1 DFQ 7

3.2 Access 7

3.3 SQL/Oracle 8

4 Technische Arten der Archivierung 9

4.1 Manuelle Archivierung 9

4.2 Automatische Archivierung im M-QIS 9

5 Ablauf manuelle Archivierung 10

5.1 Zieltyp „Archive to files“ 11

5.1.1 Archivieren oder Auslagern 11

5.2 Zieltyp „Archive to database MDB“ 15

5.2.1 Archivieren oder Auslagern 16

5.3 Zieltyp „Archive to database FD“ 16

5.3.1 Archivieren oder Auslagern 17

6 Ablauf Archivierung im Reporting - System 19

6.1 Use Case: Werte Archivieren / auslagern 20

6.1.1 Archivierung der n. – letzten Zeiteinheit 20

6.1.2 Initialarchivierungen 20

6.1.3 Kataloge in Datei speichern 21

6.1.4 Ältere Messwerte ignorieren bis 21

6.1.5 Archivierungsziel 21

6.2 Use Case: Daten löschen 23

7 Anhänge 24

7.1 Datenbank „Shrink“ oder komprimieren und reparieren 24

7.2 Löschen nicht mehr verwendeter Teile/Merkmale 24

7.3 Änderung von Grenzen zurücksetzen 26

1 Vorwort

In diesem Dokument sollen diverse Denkanstöße zum Thema der Archivierung gegeben werden. Das Thema der Archivierung erreicht die meisten Kunden leider erst, wenn der Fall schon aufgetreten ist, dass für den Tagesbetrieb ihre Datenbank zu viele Dateien hat. Das Thema sollte jedoch beim Starten eines Projekts von Anfang an berücksichtigt und eingeplant werden.

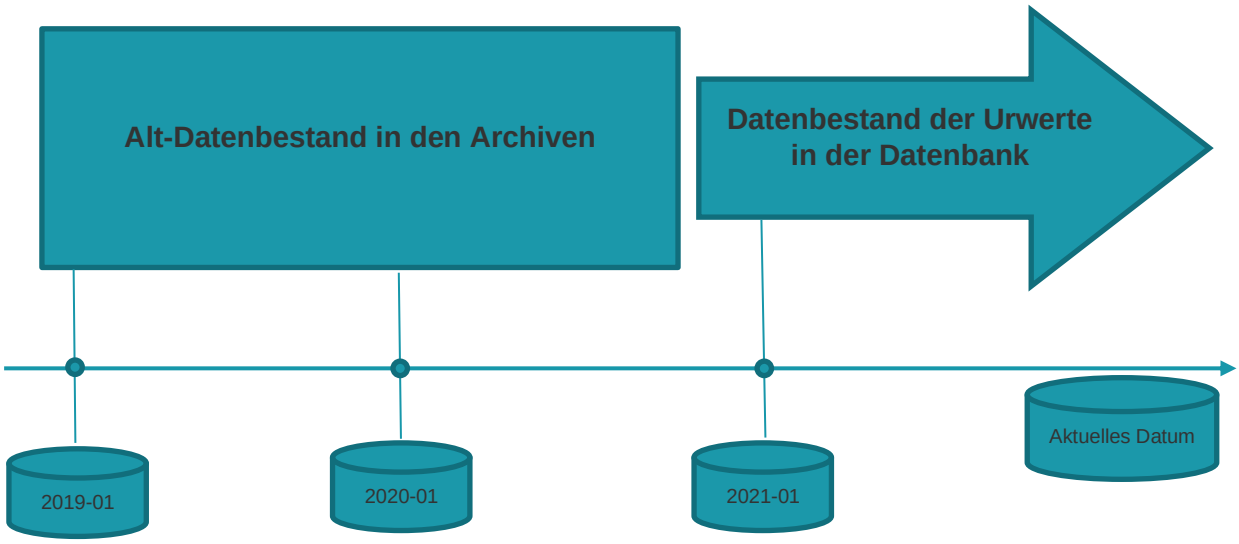
Im Rahmen von Workshops bietet hier auch das Projektteam ihre Unterstützung vor Ort an. Den Link hierzu finden Sie auf unserer Homepage.

Betrachten wir einen Zeitstrahl. Jede Firma hat die Verpflichtung, Daten aufzubewahren. Zwei bis 50 Jahre, je nach Bereich in dem eine Firma tätig ist und je nach rechtlichen Vorgaben. Jedoch müssen diese Daten nicht in der einen großen Datenbank vorhanden sein.

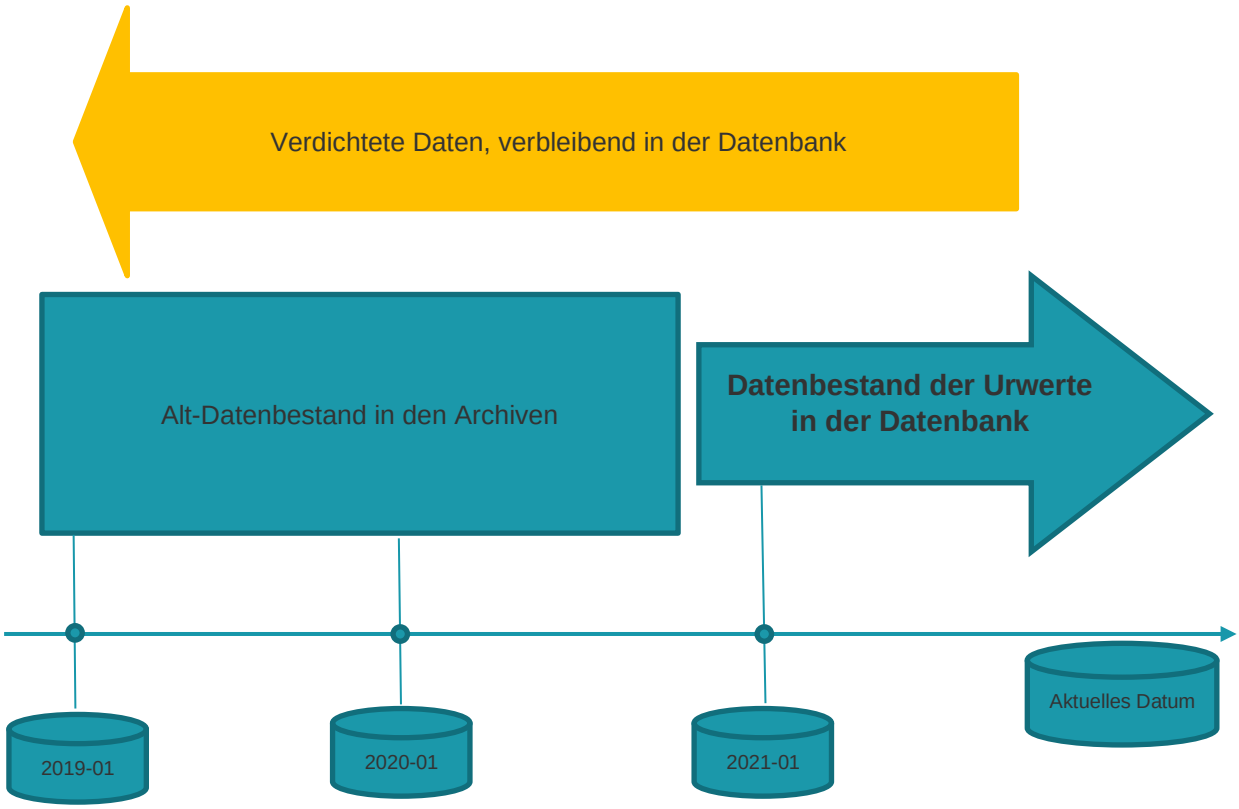
Die Überlegung lautet daher: Welcher Datenbestand muss für die aktuellen Auswertungen, für das Tagesgeschäft immer direkt in der Datenbank vorhanden sein. Aber ab welchem Alter der Daten kann es den Anwendern zugemutet werden auf Daten zuzugreifen, die in Archiven abgelegt sind.

Schauen wir uns folgendes Beispiel an:

Als Beispiel als Überlegung: mit dem aktuellen Datum **2021-10** entscheidet die Firma in der Live-DB nur noch Daten der letzten 14 Monate zu haben. Also kann alles älter als 14 Monate archiviert werden



Mit dem Produkt M-QIS können auch Daten VOR der Archivierung verdichtet werden. Durch die Verdichtung werden die berechneten Kennwerte abgespeichert, welche ohne Messwerte dann im Modul Langzeitanalyse gelesen aus ausgegeben werden können. So wäre über dieses Modul neben den Benchmarks und der zeitlichen Entwicklung der Prozesse auch ein verdichteter Blick auf ältere Daten möglich.



2 Vorüberlegungen zur Archivierung

Zu Beginn eines jeden Projektes mit der Q-DAS Software oder direkt zu Beginn der Archivierung müssen folgende Punkte bedacht werden.

		Bemerkung
Zeitbereich	Die Zeitbereichs-Überlegung umfasst 2 Punkte: • Daten bis zu welchem Zeitbereich zurück sollen in der Hauptdatenbank zur direkten Analyse behalten bleiben? • In welche Teilbereiche (Monat / Quartal / Jahr) sollen die Archiv-Pakete aufgeteilt werden?	Der Datenbestand in der Hauptdatenbank soll dem entsprechen, was im Tagesgeschäft ununterbrochen verfügbar sein muss. Die Größe der Archivteile hängt maßgeblich vom Archiv-Typ und der Datenmenge pro Einheit ab. Jeder Archiv-Typ hat eine natürliche Begrenzung was die Größe betrifft.
Archiv-Datentyp	Zur Verfügung stehen • dfq-Dateien • Access-Datenbanken • SQL/Oracle-Datenbanken	Jeder Archiv-Typ hat seine Vor- und Nachteile. In Kapitel 3 sind diese aufgelistet.
Manuell / automatisch	In jeder Windows-Applikation steht die manuelle Archivierung zur Verfügung. Im Produkt M-QIS steht eine automatische Archivierung zur Verfügung	Die manuelle Archivierung ist als Notfalllösung anzusehen, eine Empfehlung, gerade bei ständigem Datenfluss ist die automatische Auslagerung von veralteten Urwerten
Datenverdichtung vor Auslagerung	Mit dem Produkt M-QIS kann eine Datenverdichtung geschehen vor der Auslagerung	Die Datenverdichtung für das Modul Langezeitanalyse bietet die Möglichkeit die Berechnungen noch einsehen zu können, auch wenn keine Urwerte mehr dieses Zeitbereiches zur Verfügung stehen. Dies ist ausschließlich bei Verwendung von MS SQL oder Oracle – Datenbanken sinnvoll.

3 Archiv-Typen

Je nach Archivierungswerkzeug stehen mehr oder weniger Archiv-Typen zur Verfügung. Hier soll deren Vorkommen, sowie deren Vor- und Nachteile erläutert werden.

3.1 DFQ

Die Archivierung in DFQ – Dateien steht in allen Varianten zur Verfügung und ist die zu bevorzugende automatisierte Archivierungsoption. DFQ-Dateien sind das Hauptformat der Q-DAS Software und werden auch mit späteren Versionen immer lesbar sein.

Zu Bedenken beim Archivieren in DFQ-Dateien sind die Informationen, welche verloren gehen, da diese ausschließlich in Datenbanken vorgehalten werden:

- Teilemessungsstatus
- Zugeordnete Dokumente
- PDCA-Informationen
- Kataloginformationen
- Zusatzdatenspezifische Qualitätsregelkarten
- Änderungshistorie
- Mutter / Kind – Strukturen

In den meisten Fällen jedoch gibt es für diese Informationen keine Verpflichtung zur Archivierung, daher können diese vernachlässigt werden.

3.2 Access

Die Archivierung in Access-Datenbanken ist ein historisches Stück der Software. In neuen Projekten hat dies keine Bedeutung mehr.

Wenn in Access-Datenbanken archiviert wird, So kann mit der Software anschließend auf die Access-Datenbanken im reinen Lese-Modus zugegriffen werden. Das Problem ist die Datenbankgröße, Access ist auf 2GB begrenzt, daher muss die Menge für ein Archiv hier im Vorfeld überlegt und dies auch getestet werden.

Archivierung in Access-Datenbanken ist nur bei der manuellen Archivierung möglich.

3.3 SQL/Oracle

Dies ist ein Sonderfall. Da Archiv kann jedes Mal in die gleiche MS SQL/Oracle – Datenbank archiviert werden. Alle Tabellen werden mit der laufenden Archivnummer erweitert.

Auf diese Archive kann nicht zugegriffen werden über die Software. Werden die Daten benötigt, so ist ausschließlich das Wiederherstellen der Archive in die Hauptdatenbank möglich.

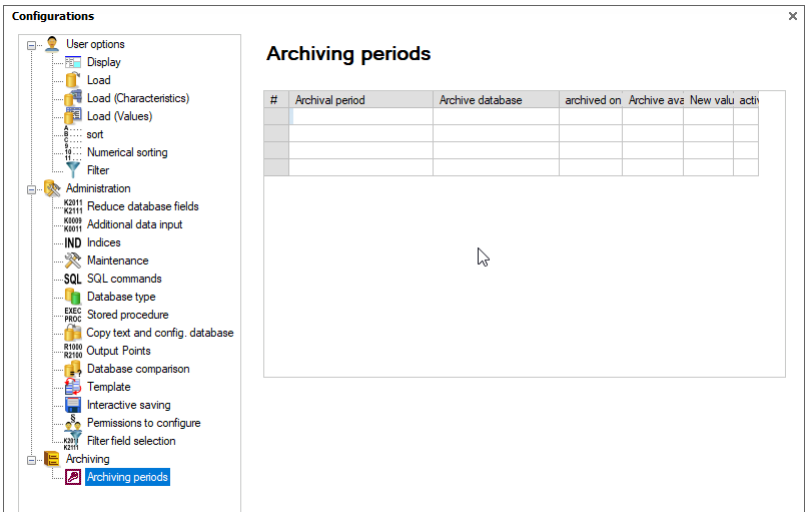
Archivierung in MS SQL / Oracle-Datenbanken ist nur bei der manuellen Archivierung möglich.

4 Technische Arten der Archivierung

Zwei verschiedene Archivierung-Werkzeuge stehen zur Verfügung. Die manuelle Archivierung (dies ist in jeder Standardapplikation möglich), sowie der automatischen Archivierung als Teil des M-QIS.

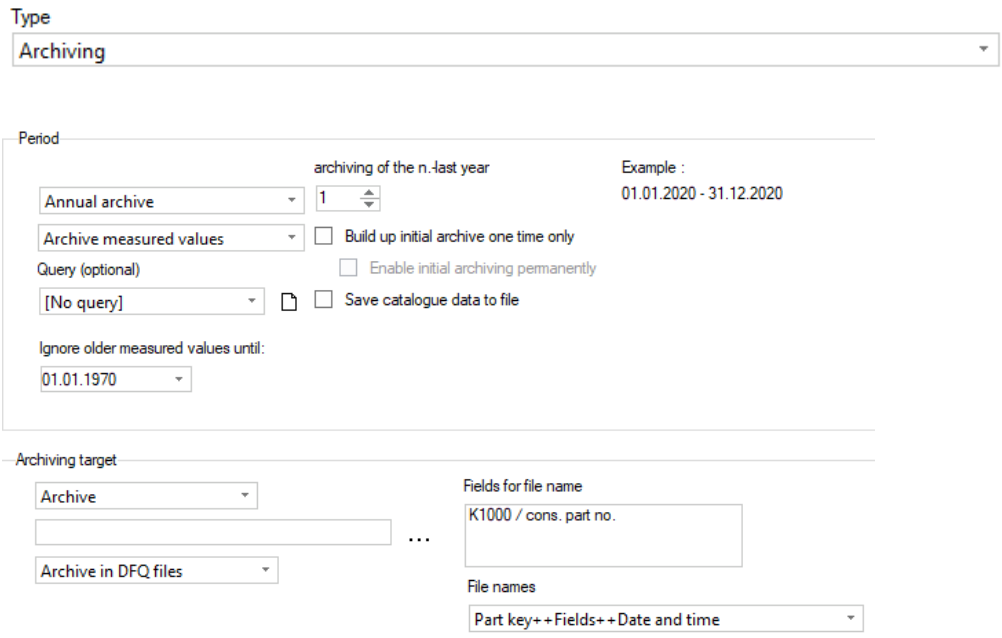
4.1 Manuelle Archivierung

Die manuelle Archivierung ist in den Datenbankoptionen aufrufbar, sofern die Berechtigung zur Datenbank-Administration gegeben ist.:



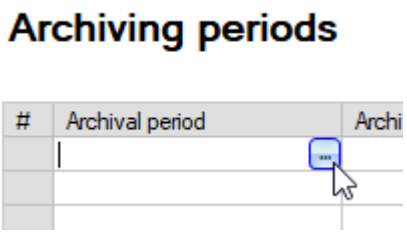
4.2 Automatische Archivierung im M-QIS

Im M-QIS Engine Reporting steht dies mit dem Typ „Archiving“ zur Verfügung

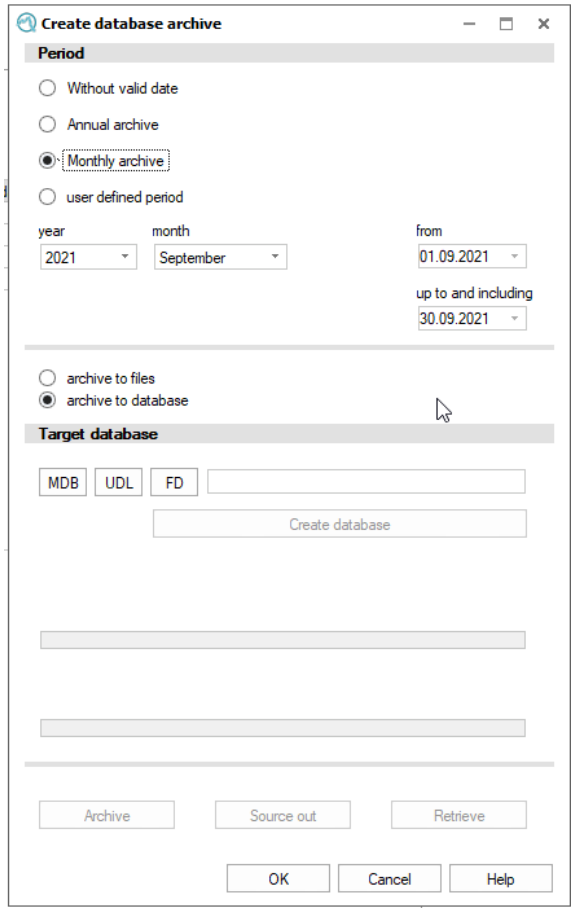


5 Ablauf manuelle Archivierung

In der Archivoption wird in die erste freie Zelle geklickt, dann kann das Menu geöffnet werden:



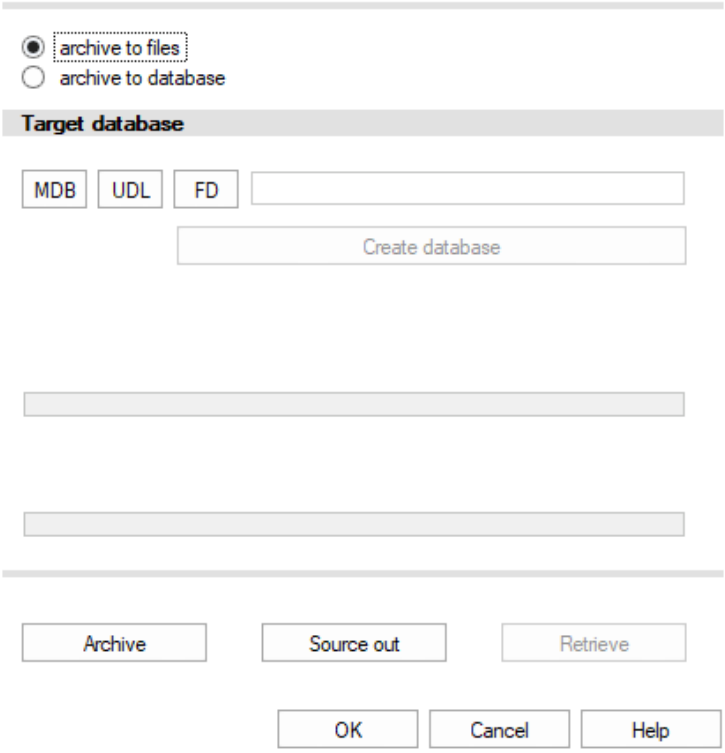
Danach öffnet sich der folgende Dialog:



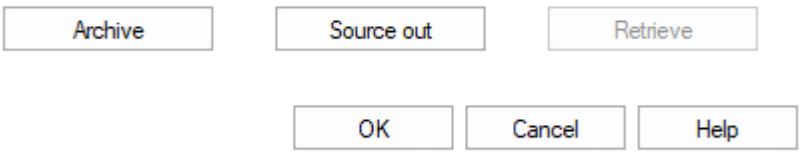
Im oberen Teil wird der Bereich des zu archivierenden Zeitraumes angegeben. Je nach Zieltyp muss hier die Größe entsprechend klein gewählt werden (Hauptsächlich, wenn das Ziel eine Access-Archivdatenbank ist.)

5.1 Zieltyp „Archive to files“

Mit dem Zieltyp „in Dateien archivieren“ hat die darunter stehende Datenbankverbindung keine Auswirkung mehr.



5.1.1 Archivieren oder Auslagern

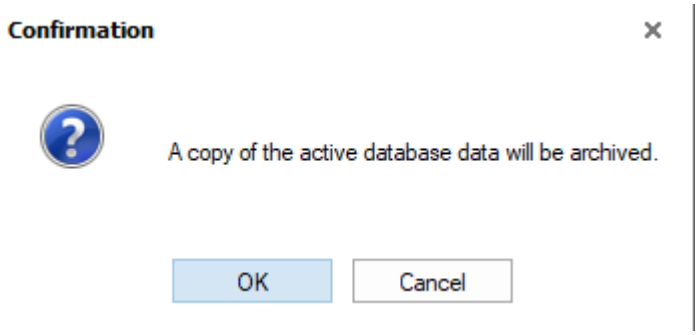


Mit den beiden Optionen „Archivieren“ oder „Auslagern“ wird begonnen das Archiv zu erstellen

Der Unterschied der beiden Optionen ist:

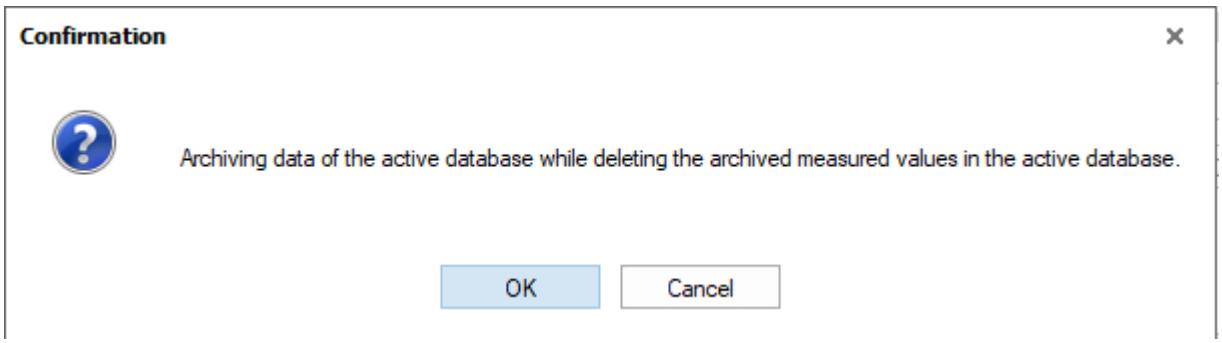
Archivieren

Es kommt ein Hinweis, dass hierdurch nur eine KOPIE erstellt wird. Es werden keine Daten ausgelagert



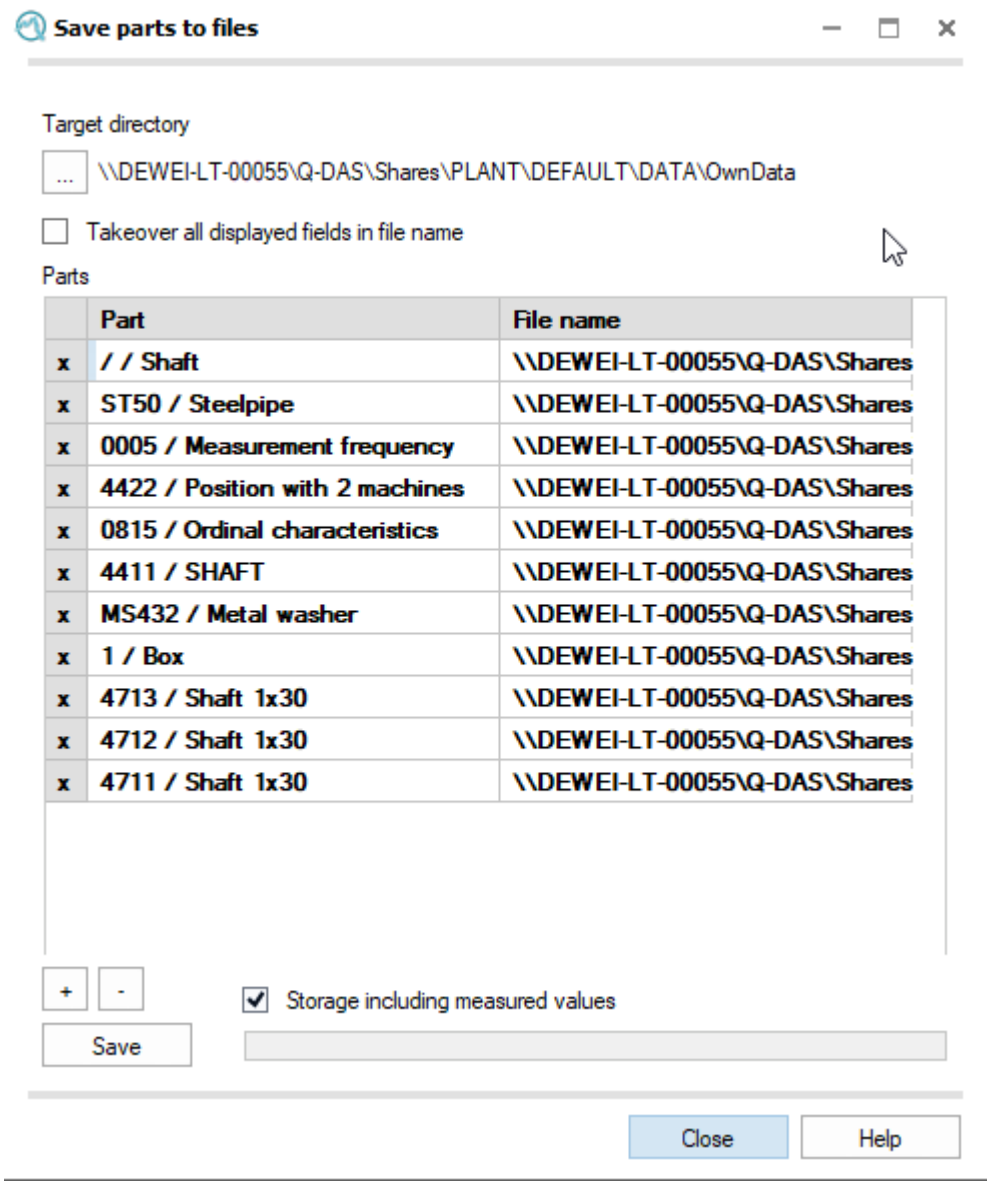
Auslagern

Es kommt ein Hinweis, dass hiermit nun auch die Messwerte, welche dann im Archiv vorhanden sind aus der Produktivdatenbank entfernt werden



Im Archiv wird für jedes Teil der Datenbank eine dfq – Datei angelegt. Ob in der dfq Werte enthalten sind, entscheidet dann nur der Zeitbereich.

Nach dem Betätigen einer der beiden Optionen öffnet sich der Dialog „Teile in Datei abspeichern“



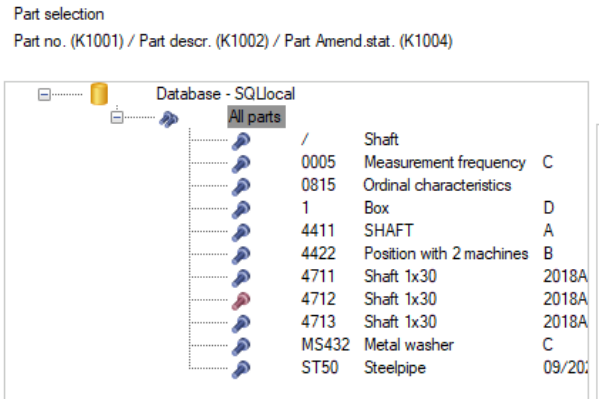
Alle Teile sind Vormarkiert. Falls gewünscht können hier einzelne Teile mit einem Klick auf das „x“ wieder deaktiviert werden.

Zielverzeichnis

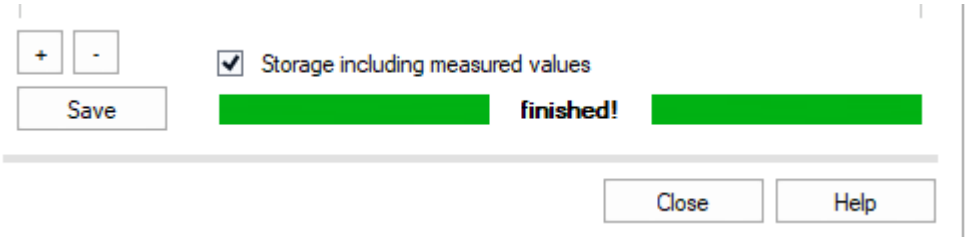
Initial wird das Standardverzeichnis zum Ablegen eigener Dateien angegeben. Optional ist ein eigenes Verzeichnis wählbar.

Alle angezeigten Felder in Dateinamen übernehmen

Im Default werden nur K1001 und K1002 in den Dateinamen zusätzlich zur internen (und damit eindeutigen) laufenden Teilenummer vergeben. Mit der besagten Option werden alle Felder, welche im Dialog „Lesen aus der Datenbank“ angegeben sind in den Dateinamen mit übernommen.



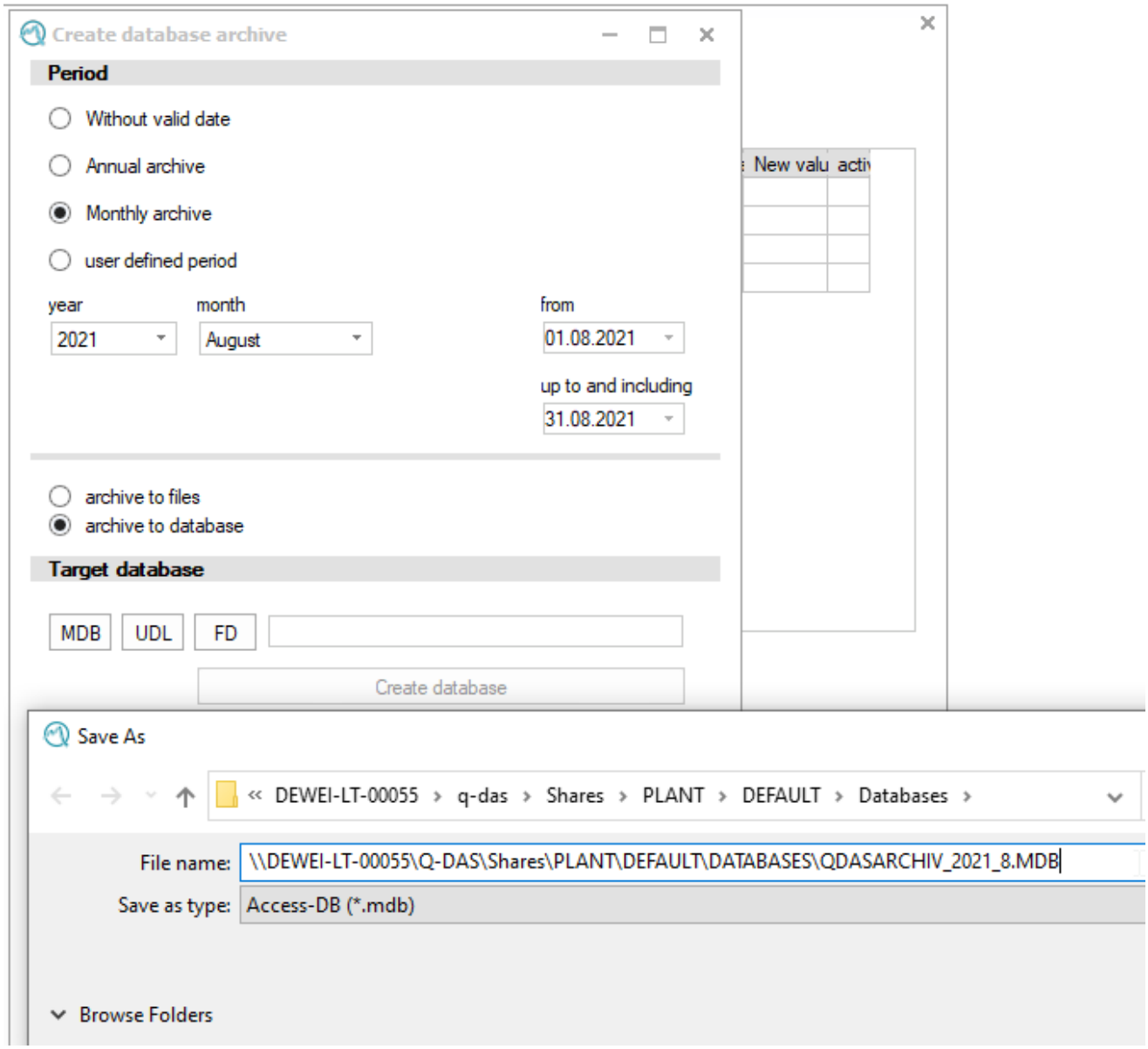
Mit einem Klick auf „Speichern“ beginnt der Archivierungs – Auslagerungsprozess. Nach dem erfolgreichen Speichern wird dies entsprechend angezeigt:



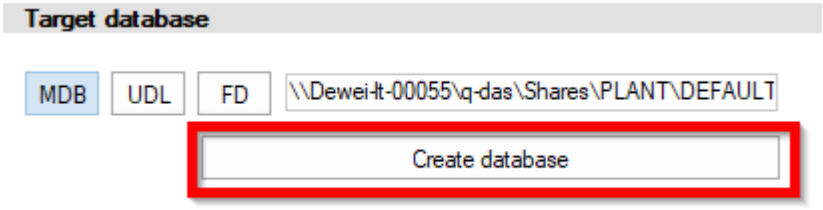
5.2 Zieltyp „Archive to database MDB“

Wird als Zieltyp „Archive to database“ gewählt, so muss dann der Datenbanktyp gewählt werden.

Wird „MDB“ gewählt, so wird der Dateiname entsprechend den Datumsvorgaben vorgeschlagen.



Anschließend muss mit „Datenbank erstellen“ diese erstellt werden:



5.2.1 Archivieren oder Auslagern

Wie auch bei DFQ-Dateien gibt es auch hier den Unterschied zwischen Archivieren (Kopie erstellen) und Auslagern (entfernen aus der Hauptdatenbank.

In der Archivdatenbank sind mit Ausnahme der Tabelle WERTEVAR (welche die Messwerte beinhält) alle anderen Tabellen eindeutige Kopien der Hauptdatenbank.

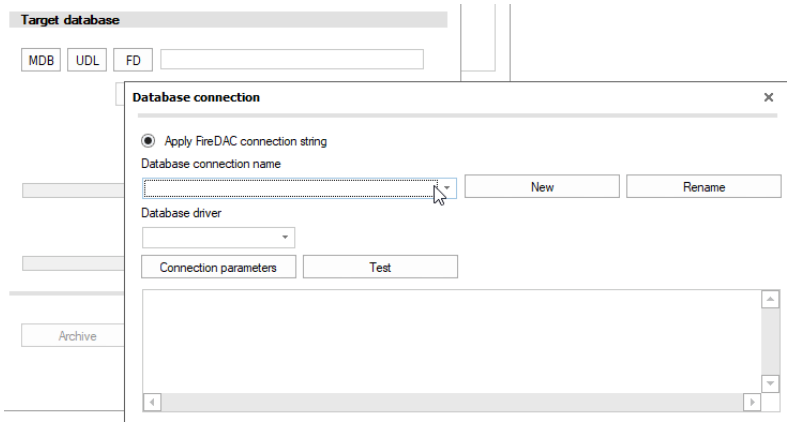
Auf die Archiv-MDB kann direkt zugegriffen werden, aber nur lesend. Die Access-DB ist schreibgeschützt.

5.3 Zieltyp „Archive to database FD“

Wird als Zieltyp „Archive to database“ gewählt, so muss dann der Datenbanktyp gewählt werden.

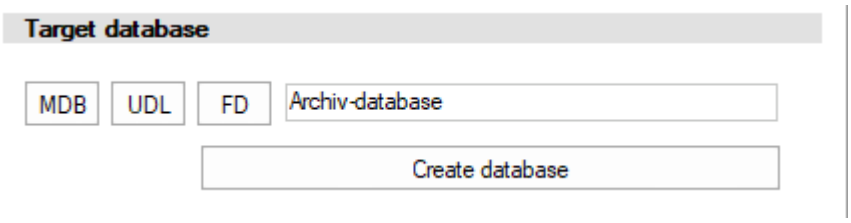
Wird „FD“ gewählt, so wird eine Firedac-Verbindung auf die SQL / Oracle – Datenbank erstellt:

Im Dialog werden keine der klassischen Datenbankverbindungen dargestellt. Es muss hier eine neue Verbindung erstellt werden:

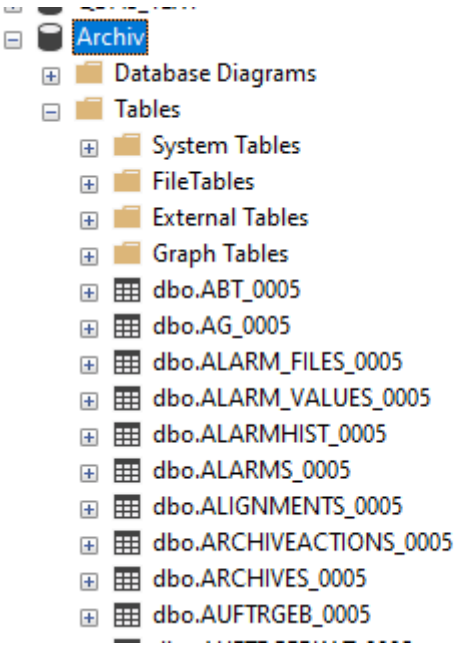


Die erste Datenbank muss eine neu angelegte und komplette leere Archiv – Datenbank sein. (keine bestehende!)

Mit „Create Database“ werden dann die Tabellen für das Archiv vorbereitet



Die Tabelle eine Archiv – Datenbank (MS SQL / Oracle) haben als Präfix die laufende Nummer des Archives. Dadurch können beliebige Archive in eine Archiv – DB ausgelagert werden.



Dadurch ist aber auch ein direkter Zugriff auf eine Archiv – Datenbank nicht möglich.

5.3.1 Archivieren oder Auslagern

Wie auch bei DFQ-Dateien gibt es auch hier den Unterschied zwischen Archivieren (Kopie erstellen) und Auslagern (entfernen aus der Hauptdatenbank.

In der Archivdatenbank sind mit Ausnahme der Tabelle WERTEVAR (welche die Messwerte beinhält) alle anderen Tabellen eindeutige Kopien der Hauptdatenbank.

Auf die Archiv-Datenbank (MS SQL / Oracle) kann NICHT direkt zugegriffen werden

6 Ablauf Archivierung im Reporting - System

Im Reporting-System wird dies als „Archiving“ – Job angelegt:

Name
Archiving

Type
Archiving

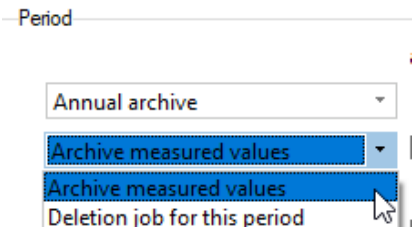
Language
0 - Default

Alle Optionen in ihrer Kombination zu beschreiben ist nicht möglich, daher wird hier auf die verschiedenen Use-Cases eingegangen

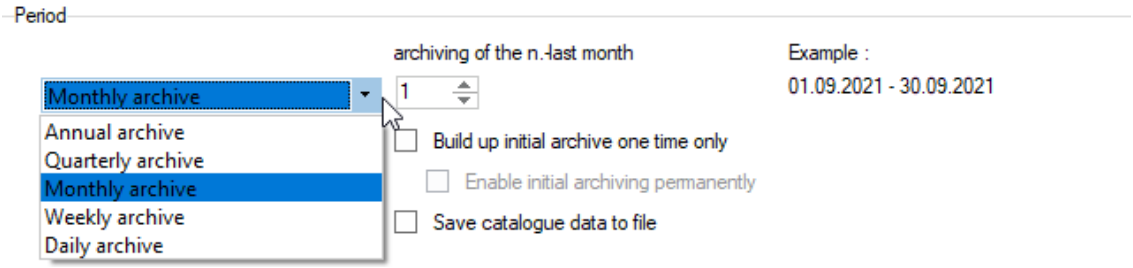
Eine Archivierung durch das M-QIS Reporting System ist ausschließlich in *.dfq – Dateien möglich.

6.1 Use Case: Werte Archivieren / auslagern

Um Messwerte zu archivieren / auszulagern wird „Messwerte archivieren“ ausgewählt:



Nun werden die Zeitbereiche definiert. Zur Auswahl stehen Tages- bis Jahres-Bereiche.



6.1.1 Archivierung der n. – letzten Zeiteinheit

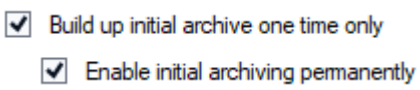
Dies ist die Angabe, welche Daten noch behalten werden sollen. Am Beispiel : Aktueller Tag ist der 13. Oktober 2021. Gewählt ist: der 3.-letzte Monat soll archiviert werden, dies wird dann auch im Beispiel rechts daneben gezeigt.

6.1.2 Initialarchivierungen

Wird begonnen mit dem M-QIS Reporting-System zu archivieren, so ist meist ein weit älterer Datenbestand vorhanden als der im Zeitbereich eigentlich angegebene. Mit der Option „Einmalige Initialarchivierung durchführen“ werden auch alle älteren Zeitbereiche in der gleichen Einheit archiviert.

Die Option ist flüchtig und würde nach dem ersten Lauf des Jobs wieder deaktiviert sein.

Wenn die Option dauerhaft gesetzt werden soll (beispielsweise, weil der Verdacht besteht, dass durch zuführende Systeme immer wieder Altdaten in die Datenbank kommen), so wird die zweite Option mit gesetzt, dann bleibt diese Option bestehen.



6.1.3 Kataloge in Datei speichern

Wenn der Verdacht besteht, dass die Kataloge sich über die Zeit dauerhaft verändern, so kann mit dieser Option jede einzelne DFQ-Datei mit dem Katalog in Datei abgespeichert werden. (K4xxx-Felder)

Beim Öffnen einer solchen Datei wechselt die Software direkt von der aktuellen Katalogeinstellung auf „Kataloge aus Datensatz“

6.1.4 Ältere Messwerte ignorieren bis

Hier kann ein Enddatum angegeben werden, bis wohin zurück Initialarchivierungen das Datum prüfen sollen.

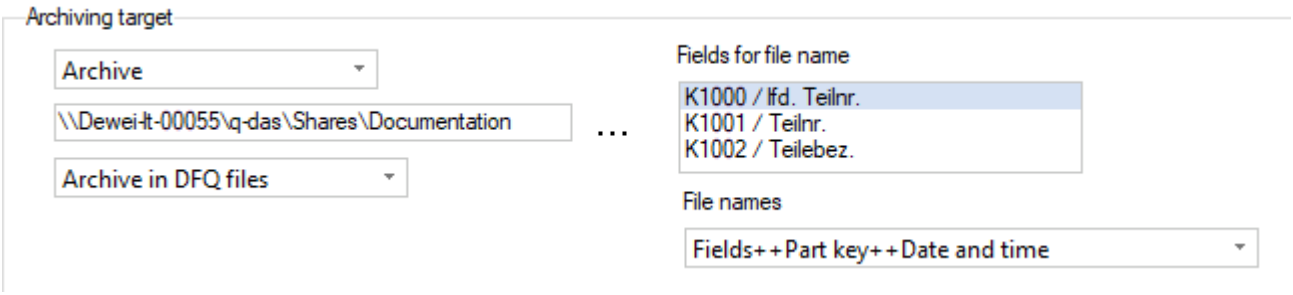
Der Grund für die Existenz dieser Option ist folgender: Wenn messwertgenerierende Systeme Datum und Zeit bei Werten nicht oder falsch geschrieben hatten, und der Upload nicht darauf eingestellt war dies zu korrigieren, so können Werte vom 01.01.1970 existieren, der 0-Zeit. Um hier keine Monatsarchive ohne Inhalt von 1970 bis ins aktuelle Jahr zu bekommen kann hier ein Maximaldatum gesetzt werden.

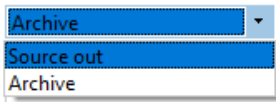
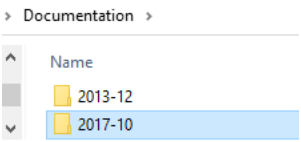
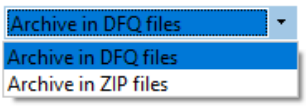
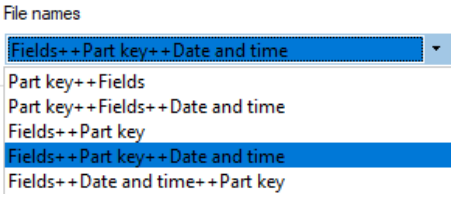
In einem sauber aufgesetzten System sollte diese Option ohne Bedeutung sein.

6.1.5 Archivierungsziel

In dieser Sektion wird angegeben in welcher Form und wohin die Dateien platziert werden.

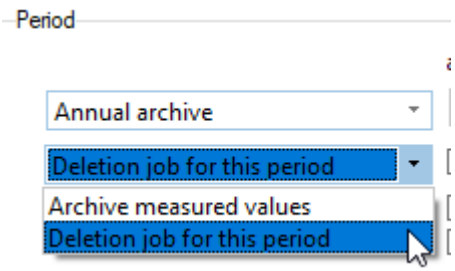
Neben den Grundsätzlichen Einstellungen wie auch bei der manuellen Archivierung sind hier weitere Angaben möglich, welche im Folgende als Tabelle erläutert werden.



Archivieren / Auslagern  	Wie auch bei der manuellen Archivierung kann hier zwischen Archiv (Kopie) und dem effektiven Auslagern entschieden werden
Ziel	Hier ist ein Hauptordner auszuwählen. In diesem werden, abhängig von der gewählten Zeiteinheit Unterordner erstellt. Beispiel Monatsarchive 
Dateityp  	Aufgrund der hohen Kompressionsrate von dfq – Dateien empfiehlt es sich, als Dateityp gezippte dfq – Dateien abzulegen. Mit der Software kann mit der Option „Öffnen aus zip“ auch direkt eine gezippte dfq – Datei geöffnet werden.
Feldnamen	Hier können Felder zur Dateinamenserzeugung gewählt werden.
Dateinamensgenerierung	Der Dateiname selbst enthält immer den Partkey, den internen Schlüssel, da nur dieser eindeutig ist in der Datenbank. Mit der Auswahl hier kann die Reihenfolge der Namen bestimmt werden. Datum/Zeit sind die aktuelle Zeit, also wann die Datei erzeugt wurde. 

6.2 Use Case: Daten löschen

Um „Daten“ zu löschen wird „Lösch-Job für diesen Zeitraum“ ausgewählt:



Hier handelt es sich um eine **nicht rückgängig zu machende Handlung!** Nur in Absprache mit der lokalen IT und zyklischen Sicherungen der Datenbank sollte diese Option genutzt werden!

„Daten“ sind im Allgemeinen:

- Messwerte
- Gespeicherte Ergebnisse
- Alarme
- Signaturen

Die Einstellung des Zeitbereiches ist analog dem Archivieren / Auslagern.

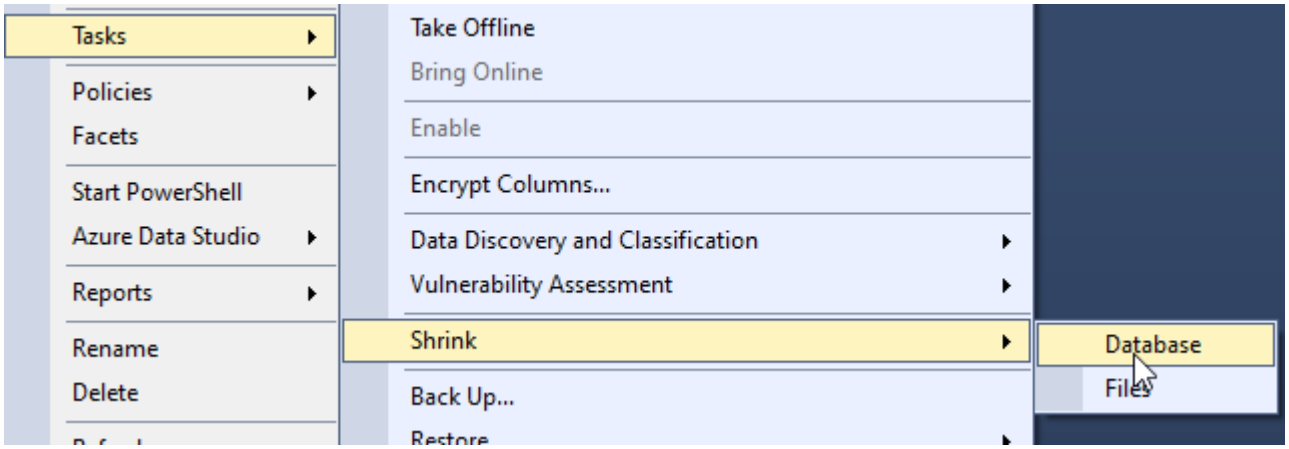
Messwerte löschen	Analog des Auslagerns, werden hier die Messwerte nur gelöscht ohne diese vorher zu sichern
Ergebnisse löschen	Gespeicherte Ergebnisse sind die komprimierten Daten der Langzeitanalyse, welche mit dem Reporting – System verdichtet wurden (Aufteilungszuordnung für das Abspeichern von Ergebnissen). Es können auch nur Daten einer bestimmten Aufteilungszuordnung gewählt werden
Alarme löschen	Alarme werden entweder über den Upload (Einzelwertalarme) oder über das procella in der Datenbank abgespeichert.
Signaturen löschen	Signaturen werden z.B. über die Berichtsansicht oder über Automatismen in procella in der Datenbank abgelegt.

7 Anhänge

Wenn aufgrund überlaufender Datenbanken das erste Mal Daten ausgelagert/gelöscht wurden, empfiehlt es sich über diverse weitere „Bereinigungen nachzudenken“

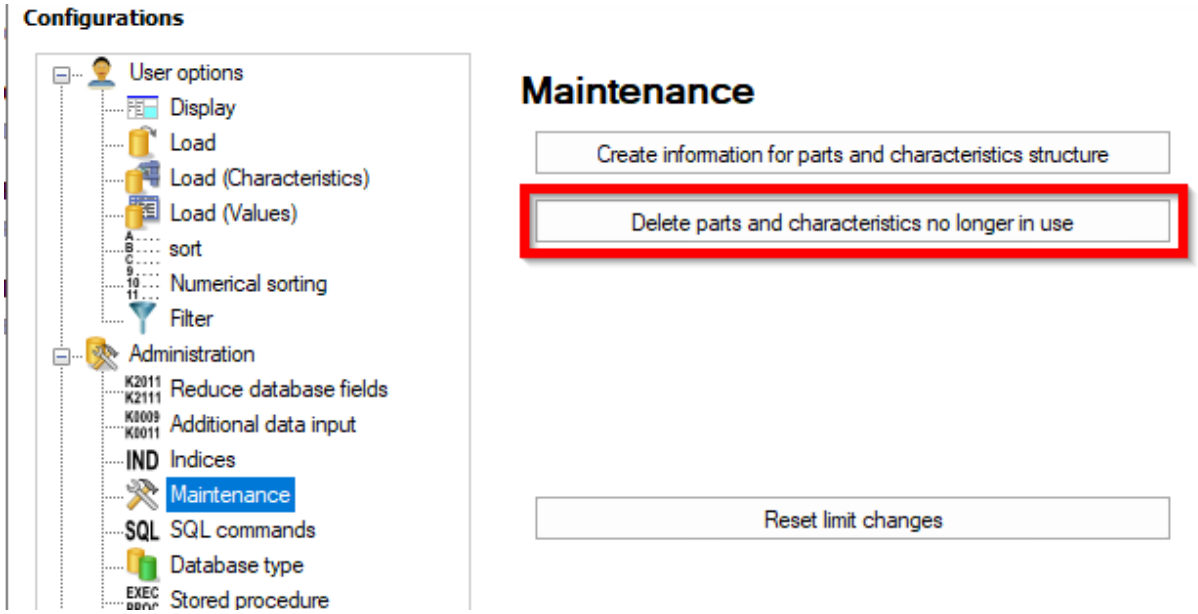
7.1 Datenbank „Shrink“ oder komprimieren und reparieren

Nach größeren Bereinigungen sollten die Datenbanken „komprimiert“ werden. Im Fall von MS SQL-Datenbanken wäre dies beispielsweise der Task „Skrink“ mittels SQL Server Management Studio.

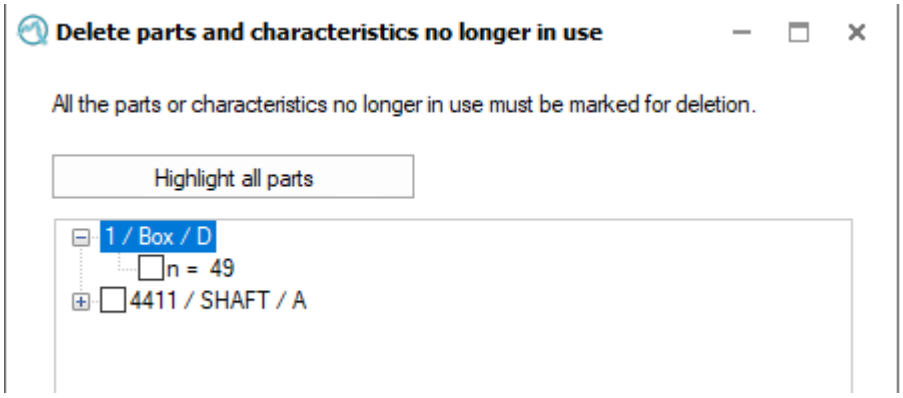


7.2 Löschen nicht mehr verwendeter Teile/Merkmale

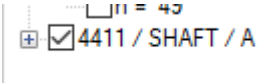
Diese Frage kommt nach dem manuellen Archivieren automatisch auf, kann aber über die Datenbankoptionen jederzeit gestartet werden



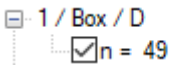
Nach dem Betätigen der Schaltfläche und der Suche wird folgender Dialog angezeigt:



Dargestellt werden alle Teile, welche gar keine Messwerte mehr haben in allen Merkmalen, direkt zu markieren VOR dem Teil:



Wenn bei einen Teil nur bestimmte Merkmale leer sind, könnten die Merkmale noch einzeln markiert werden, nicht jedoch das gesamte Teil:



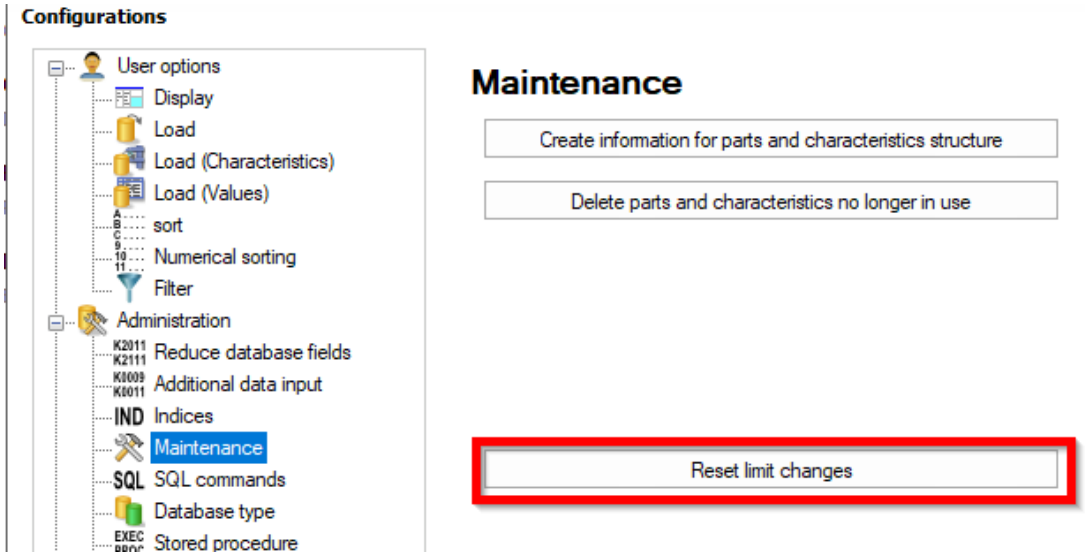
Die Option „Alle Teile markieren“ würde ausschließlich komplett leere Teile markieren.

Wenn von Datenbank in Datenbanken archiviert wurde, so existieren eventuell keine leeren Teile! Wenn von einem Teil der letzte Messwert archiviert wurde, wird ein „Platzhalter-Messwert“ geschrieben, der über die Software selbst nicht gesehen werden kann. Dies dient dazu, dass die Option „leere Teile finden“ diese Teile NICHT findet, da für eventuelle Wiederherstellungen die interne laufende Teilenummer behalten werden muss.

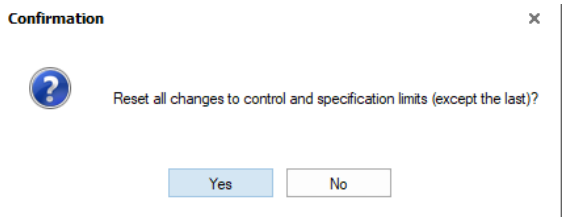
Dies sollte immer erst nach einer vorherigen Sicherung der Datenbank durch die lokale IT geschehen!

7.3 Änderung von Grenzen zurücksetzen

Auf gleicher Registerkarte befindet sich die Option „Änderung von Grenzen zurücksetzen“



Kurze Erläuterung: Auch ohne die aktivierte Änderungshistorie wird jede Änderung an Spezifikationsgrenzen oder Eingriffsgrenzen abgespeichert. Mit dieser Option kann dies für alle Datensätze der Datenbank bereinigt werden, so dass nur noch der letzte Stand (was die Software auch direkt lesen würde) existent ist.



Eine Detailauswahl ist hier nicht möglich!

Dies sollte immer nur nach eine Sicherung der Datenbank durch die lokale IT geschehen!