QS-DocumentAssembler

Version 4.6

DocumentAssembler

By QS QualitySoft GmbH

Produktdokumentation

Software-Version 4.6

01/2024

© 2003-2024 QS QualitySoft GmbH

QS QualitySoft GmbH
Tempowerkring 21a
21079 Hamburg
Tel. +49 (0)40 790 100 40
Fax +49 (0)40 790 100 44

info@qualitysoft.de

https://www.qualitysoft.de

Inhalt

1 Einführung ... 5
1.1 Aufgabenstellung ... 5
1.2 Funktionsüberblick QS-DocumentAssembler ... 5
1.2.1 Unterstützte Barcodetypen ... 6
1.2.2 Verarbeitung von Dokumenten ... 6
1.3 Standard und Professional Edition ... 8
2 Voraussetzungen / Installation ... 9
2.1 Systemvoraussetzungen ... 9
2.2 Installation ... 9
2.3 Demo-Version ... 10
2.4 Vollversion ... 10
3 Bedienung des QS-DocumentAssembler ... 11
3.1 Hauptfenster ... 11
3.1.1 Jobliste ... 11
3.1.2 Ausführung von Jobs ... 12
3.1.2.1 Manuelle Ausführung ... 12
3.1.2.2 Scheduler ... 12
3.1.3 Verarbeitung ... 13
3.1.4 Die Menüleiste ... 13
3.1.5 Die Erstellung eines neuen Jobs ... 15
3.1.6 Die Bearbeitung von Jobs ... 15
3.2 Einstellungen ... 16
3.2.1 Registerkarte "Job allgemein" ... 16
3.2.2 Registerkarte "Datei/Verzeichnis" ... 24
3.2.2.1 Die Konfiguration von Outlook365 zum Senden und Abrufen von Mails ... 32
3.2.3 Registerkarte "Verarbeitung" ... 43
3.2.3.1 Verarbeitung Einzelseitenmodus (BLAU) ... 44
3.2.3.2 Verarbeitung Stapelmodus (GRÜN) ... 47
3.2.3.3 Die Benennung der Zieldateien in der Verarbeitung von Einzelseiten- und Stapelmodus ... 51
3.2.3.4 ZUGFeRD-Verarbeitung ... 52
3.2.3.5 Konvertierung des Dateiformates ... 54
3.2.4 Registerkarte "Zieldateien" ... 56

3.2.5 Registerkarte "Barcode/OCR" ... 65
3.2.5.1 Barcode-Erkennung ... 65
3.2.5.2 OCR-Erkennung ... 67
3.2.5.3 Filter ... 67
3.2.5.4 Barcodeeinstellungen ... 72
3.2.5.5 OCR-Einstellungen ... 80
3.2.6 Registerkarte "Imprinter" ... 86
4 Plausibilitätsprüfung ... 89
5 Einbindung des QS-DocumentAssembler in die Batchverarbeitung ... 91
6 QS-DocumentAssembler Dienst ... 92
6.1 Dienst-Administration ... 92
6.1.1 Dienst-Administration über den QS-DocumentAssembler ... 92
6.1.2 Dienst-Administration über den Windows Dienstmanager ... 93
6.1.3 Dienst-Administration über die Kommandozeile ... 94
6.2 Dienst-Konfiguration ... 94
7 Anhang ... 95
7.1 Protokolldatei ... 95
7.2 Unterstützte Dateiformate ... 96
7.3 Hinweise zum Adobe PDF-Dateiformat ... 97

Redaktionelle Anmerkung

QualitySoft bemüht sich, diese Dokumentation immer aktuell zu halten. Hinweise auf Fehler oder unverständliche Stellen nehmen wir gern entgegen. Der QS-DocumentAssembler wird stetig gepflegt, verbessert und erweitert. Änderungen am Produkt sind so jederzeit möglich.

QS-DocumentAssembler, QS-DocumentAssembler.NET und QS-Barcode sind urheberrechtlich geschützt durch © QS QualitySoft GmbH.

Die von QS-DocumentAssembler genutzte QS-Barcode SDK zur Erkennung von Barcodes verwendet die Open Source FreeImage Library zum Lesen der Bilddateien.

Den Source Code und Details zur FreeImage Public License 1.0 finden Sie im Internet unter http://freeimage.sourceforge.net/license.html.

Microsoft, Microsoft .NET Framework sowie Microsoft Windows sind Warenzeichen oder eingetragene Warenzeichen der Microsoft Corporation.

Weitere in diesem Dokument erwähnte Software- und Hardwarebezeichnungen sind in den meisten Fällen auch eingetragene Marken und unterliegen als solche den gesetzlichen Bestimmungen.

1 Einführung

1.1 Aufgabenstellung

Mit Dokumentenscannern gescannte, mit digitalen Kameras erfasste oder von Faxservern empfangene Dokumente liegen als Bilddateien vor. Auf diesen Bilddateien sollen Barcodes erkannt werden. Ebenso können Adobe PDF-Dokumente Barcodes enthalten, die gelesen werden sollen. Die erkannten Barcodes sollen für die weitere Verarbeitung bereitgestellt oder Dateien sollen nach erkannten Barcodes umbenannt werden.

1.2 Funktionsüberblick QS-DocumentAssembler

Der QS-DocumentAssembler ist eine leistungsfähige, flexibel konfigurierbare Software zur massenweisen Lesung von Barcodes und OCR-Bereichen aus Bildern oder Adobe PDF-Dateien.

Die automatisierte Erkennung sowie die Sortierung, Indizierung, Strukturierung, Zuordnung und Umbenennung von gescannten Dokumenten anhand von Barcodes sind die häufigsten Einsatzgebiete.

Es können auf der ganzen Seite Barcodes unter beliebigen Winkeln gesucht werden. Auch Dateiformate mit mehreren Seiten pro Datei werden unterstützt. Von beliebig vielen Barcodes können die Ergebnisse in Dateien geschrieben werden.

QS-DocumentAssembler kann sowohl 1-dimensionale (lineare) Barcodes als auch 2-dimensionale Barcodes der Typen PDF417, DataMatrix, QR-Code und Aztec-Code lesen, die sehr viele Daten enthalten können. So werden in einer Version für die Finanzverwaltung ("Elster-Barcode") mehr als 800 verschlüsselte Zeichen erkannt und für die Folgeverarbeitung zur Verfügung gestellt. Die Erkennungszeit für einen PDF417 Barcode, wie er im Bankenbereich eingesetzt wird, liegt bei unter 50 ms.

Die Barcode-Erkennung der Anwendung basiert auf der "QS-Barcode Library", die auch als eigenständiges Produkt (QS-Barcode SDK) für Software-Entwickler verfügbar ist.

Die OCR-Erkennung auf Dokumenten kann so konfiguriert werden, dass die Ergebnisse als Barcodeergebnis zur weiteren Verarbeitung der Dokumente verwendet werden kann.

Der QS-DocumentAssembler unterstützt die Erstellung und Verwaltung mehrerer individueller Konfigurationen, im folgenden Job genannt. In jedem Job wird ein individuelles Set von Verarbeitungsmerkmalen zusammengefasst. Hierzu gehören z.B. Quell- und Zielverzeichnis, zu erkennende Barcodetypen, Verarbeitung der Zieldateien und vieles mehr.

Die Jobs des QS-DocumentAssembler können manuell ausgeführt oder so konfiguriert werden, dass Sie über einen Scheduler in definierten Zeitabständen einmalig oder kontinuierlich ausgeführt werden.

Der QS-DocumentAssembler ist auch für den Batch-Betrieb konzipiert, bei dem automatisch massenhaft Belege verarbeitet werden.

Der QS-DocumentAssembler bietet Ihnen weiterhin die Möglichkeit, die Verarbeitung von Dokumenten von lokalen Laufwerken und Netzwerklaufwerken ohne Benutzereingriff auch als Windows Dienst durchzuführen.

1.2.1 Unterstützte Barcodetypen

1-dimensionale Barcodetypen (Lineare Barcodes)

- Code39
- Code93
- Code 32
- Code 11
- Codabar
- Code 128
- Codablock F
- EAN 8 und 13
- EAN 128 / UCC 128
- UPC A / UPC E
- Code 2/5 (Interleaved, Industrie, IATA, BCD-Matrix, Datalogic)
- GS1-Familie (EAN-13 und DataBar)
- Pharmacode
- Patchcode

2-Dimensionale Barcodetypen

- Aztec
- Data Matrix
- QR Code
- PDF417.

1.2.2 Verarbeitung von Dokumenten

Im Rahmen der Verarbeitung von Dokumenten werden im QS-DocumentAssembler die folgenden Anwendungsfälle unterschieden.

Auf jeder Seite eines Quelldokumentes ist ein Barcode aufgebracht

Bei den Quelldateien kann es sich sowohl um ein- als auch um mehrseitige Dokumente handeln. Wenn ein Barcode gefunden wurde, wird die Seite nach den eingestellten Regeln im definierten Zielverzeichnis abgelegt. Seiten ohne Barcodes werden im definierten Fehlerverzeichnis abgelegt.

Zusammenhalten eines mehrseitigen Quelldokumentes

Es wird auf allen Seiten des Quelldokumentes nach Barcodes gesucht. Die Benennung des gesamten Dokumentes erfolgt nach dem ersten gefundenen Barcode. Wird kein Barcode gefunden, wird das Dokument in ein definiertes Fehlerverzeichnis verschoben.

Jeder Barcodewechsel definiert den Beginn eines neuen Stapels (Zieldatei)

Bei den Quelldateien kann es sich sowohl um ein- als auch um mehrseitige Dokumente handeln. Dabei ist sicher zu stellen, dass die gescannten Dokumente fortlaufend nummeriert sind.

Wenn eine Seite mit Barcode gefunden wird, wird ein neuer Stapel (Zieldatei) erstellt. Alle folgenden Seiten ohne Barcode werden an diesen Stapel angehängt bis ein neuer vom vor-

herigen Barcode abweichender Barcode gefunden wird, der wiederum einen neuen Stapel beginnt.

Jeder gefundene Barcode definiert den Beginn eines neuen Stapels (Zieldatei)

Bei dieser Konstellation wird, abweichend von der vorherigen Definition, bei jedem gefundenen Barcode ein neuer Stapel definiert. Dies dient u.a. dazu, im Stapelmodus Dokumente mit identischen Barcodes zu trennen anstatt sie in eine Zieldatei zu schreiben.

Barcode definiert den Beginn eines neuen Stapels mit Entfernung von Deckblättern/Trennblättern.

Sie erstellen Trennblätter mit Barcode, die Sie vor dem Scannen vor Ihren Dokumentenstapel legen.

Bei den Quelldateien kann es sich sowohl um ein- als auch um mehrseitige Dokumente handeln. Dabei ist sicher zu stellen, dass die gescannten Dokumente fortlaufend nummeriert sind. Der Stapel (Zieldatei) wird durch den Barcode erzeugt, jedoch werden Seiten mit einem Barcode entfernt und nicht in das Zieldokument übernommen.

Das Zieldokument bekommt den Namen des Barcodes des vorhergehenden Deckblattes. Diese Option bietet sich an, wenn Ihre Dokumente keine Barcodes aufweisen.

Barcode definiert den Beginn eines neuen Stapels und Folgeseiten haben einen fest definierten Barcode

Es besteht immer die Gefahr, dass Barcodes nicht erkannt werden und dann zwei Stapel zu einem zusammengefasst werden. Wenn man aber auf jeder Folgeseite einen Platzhalter-Barcode mit einem festen, gleichbleibenden Wert anbringt, kann das fehlerhafte Zusammenfassen von zwei Stapeln verhindert werden.

Identischer Barcode definiert Stapelbeginn und –ende.

Eine weitere Möglichkeit, das fehlerhafte Zusammenfassen von zwei Stapeln zu verhindern besteht darin, den Beginn und das Ende eines Stapels mit einem Identischen Barcode zu kennzeichnen. Wenn der zweite Barcode hier nicht gefunden werden kann, wird der gesamte Bereich bis zum nächsten gefundenen Barcode als Dokument im Fehlerverzeichnis gespeichert. Auch hier besteht die Möglichkeit, die Seiten mit den Barcodes zu entfernen.

Auslesen und Protokollieren von Barcodes in Dokumenten

QS-DocumentAssembler kann auch benutzt werden, um viele Barcodes zu Protokollzwecken aus einer Quelldatei zu lesen und in einer Ergebnisdatei abzulegen. Diese Ergebnisdatei kann nach der Erkennung von Drittprogrammen weiterverwendet werden.

Anwendungsbeispiele sind Protokollierungen für den Warenausgang, bei dem viele kleine Barcode-Etiketten mit den Seriennummern auf große Kontrollbögen geklebt werden, die dann gescannt werden. Die Lesung mit QS-DocumentAssembler ist hierbei einfacher und schneller als mit Handlesern.

Im Rahmen der hier beschriebenen Verarbeitungsmöglichkeiten bietet der DocumentAssembler weitere Funktionalitäten, die je nach Bedarf verwendet werden können.

Grundsätzlich besteht die Möglichkeit in allen Anwendungsfällen, den Dateinamen der Zieldatei über den variablen Zieldateinamen individuell anzupassen. Hierfür steht eine Vielzahl vordefinierter Variablen zur individuellen Gestaltung des Zieldateinamens zur Verfügung.

Des Weiteren können in allen Anwendungsfällen Leerseiten aus den Quelldokumenten herausgefiltert werden. Hier besteht die Möglichkeit, über die Definition eines Schwärzungsgrades einer Leerseite diese herauszufiltern.

Zusätzlich besteht die Möglichkeit, das gleiche Set an Variablen zum individuellen Aufbringen eines Imprinter-Stempels auf eine frei definierbare Position im Zieldokument zu benutzen.

Falls sich mehrere Barcodes auf den Quelldokumenten befinden, können Sie den gewünschten Barcode über Filter mit Regulären Ausdrücken herausfiltern.

Gelegentlich kommt es vor, dass Seiten verkehrt herum eingescannt werden. Hier besteht die Möglichkeit, diese Seiten in Richtung des Barcodes zu drehen.

Beim Schreiben der Zieldateien können Sie auswählen, welches Format die Zieldatei haben soll. Wir unterstützen die Formate PDF und TIFF als ein- und mehrseitiges Format sowie die Formate BMP, PNG, JPG, GIF als einseitiges Format. Falls Sie beim Scannen also ein Grafikformat produzieren müssen, als Zielformat aber PDF benötigen, können Sie das Ausgabeformat problemlos wechseln.

Hinweis

Die Barcode-Erkennung basiert auf der QS-Barcode Library.

Die optimalen Barcode-Einstellungen können mit dem Freeware-Programm bcTester gefunden werden: https://www.bctester.de

Zur detaillierten Erläuterung der einzelnen Parameter sei auf die entsprechende Dokumentation des QS-Barcode SDK verwiesen, zu finden im Internet:

https://qualitysoft.de/wp-content/uploads/qsbarsdk.pdf

Basisinformationen zur QS-Barcode Erkennung (Typübersicht, Druckhinweise, etc.)

https://qualitysoft.de/wp-content/uploads/bcinf.pdf

QS-DocumentAssembler Preisliste und Bestellfax:

https://qualitysoft.de/wp-content/uploads/docasm_preise.pdf

Wenn Probleme bei der Erkennung Ihrer Barcodes auftreten, können Sie jederzeit auf uns zukommen.

Senden Sie einfach Ihre Bilder an info@qualitysoft.de!

1.3 Standard und Professional Edition

Der QS-DocumentAssembler liegt in einer **Standard**- und einer **Professional Edition** vor. Die Professional Edition unterscheidet sich von der Standard Edition dadurch, dass die Professional Edition die Verarbeitung von Dokumenten als Windows Dienst sowie eine beliebige Anzahl von parallelen Prozessen bei der gleichzeitigen Verarbeitung mehrerer Jobs unterstützt.

2 Voraussetzungen / Installation

2.1 Systemvoraussetzungen

QS-DocumentAssembler ist lauffähig unter den Microsoft Windows Betriebssystemen Server 2008, Server 2012, Server 2016, Server 2019, Server 2022, Windows 7/8/10/11. Das Microsoft .NET Framework 4.8 muss installiert sein. Jeder PC mit mindestens 210 MB freiem Festplattenspeicher eignet sich für den Einsatz des **QS-DocumentAssembler**. Die Bilddateien, die mit der Software verarbeitet werden sollen, nehmen aber oft ein Vielfaches dieser Speichermenge in Anspruch. Die Angabe 210 MB bezieht sich auf die Software allein.

Zum Scannen mit Dokumentenscannern und zur automatischen Ablage der digitalisierten Bilder, die **QS-DocumentAssembler** verarbeiten soll, steht eine große Auswahl von Geräten zur Verfügung, die mit unterschiedlicher Scan-Software betrieben werden können.

QualitySoft berät Sie gern bei der Auswahl entsprechender Komponenten und verfügt auch selbst über ein entsprechendes Angebot an Produkten.

2.2 Installation

Die aktuelle Evaluationsversion des QS DocumentAssembler steht unter https://www.qualitysoft.de/produkte/qs-documentassembler zur Verfügung.

Der QS-DocumentAssembler unterstützt nur 64-Bit Betriebssysteme.

Mit dem Installationsprogramm des QS-DocumentAssembler kann jetzt über bestehende Installationen ab Version 4.0 eine Aktualisierung durchgeführt werden. Eine vorherige Deinstallation der alten Version ist nicht mehr nötig.

Die Evaluationsversion läuft im Demomodus. Vor dem Setup des QS DocumentAssembler müssen die Installationsdateien aus dem entsprechenden Archiv in ein beliebiges Verzeichnis entpackt werden.

Zur Installation gehen Sie bitte wie folgt vor:

1. Installieren Sie ggf. das Microsoft .NET Framework 4.8, falls es auf dem Zielsystem nicht vorhanden ist.
2. Kopieren Sie ggf. Ihre Lizenzdatei (qsbc.lic) für die Vollversion in das Verzeichnis, in das Sie das Archiv mit der Demo-Version entpackt haben und überschreiben die bestehende Demo-Lizenz.
3. Starten Sie die Anwendung SETUP.EXE
4. Folgen Sie den Anweisungen des Setup-Programms, ändern Sie, wenn gewünscht, das Zielverzeichnis der Installation
5. Während der Ausführung von SETUP.EXE wird automatisch auch die individuell für Sie erstellte Lizenzdatei QSBC.LIC aus dem Installationsverzeichnis in das Zielverzeichnis der Installation kopiert (entfällt in der Evaluationsversion).
6. Sie können den QS-DocumentAssembler nun über das Startmenü starten, eine Programmgruppe wurde dort eingerichtet.
7. Konfigurieren Sie den QS-DocumentAssembler Ihren Anforderungen entsprechend.

Hinweise:

- Wenn sie den QS-DocumentAssembler Dienst nutzen wollen empfiehlt es sich, die Voreinstellung zur Installation für alle Benutzer nicht zu ändern. Andernfalls kann der QS-DocumentAssembler Dienst nur für den während der Installation angemeldeten Benutzer ausgeführt werden, da die notwendige Registrierung von Komponenten für andere Benutzer fehlt.

Wenn Sie mit einer Demo-Lizenz arbeiten oder die Lizenzdatei vom QS DocumentAssembler nicht gefunden wird, schaltet die Anwendung automatisch in den Demomodus um.

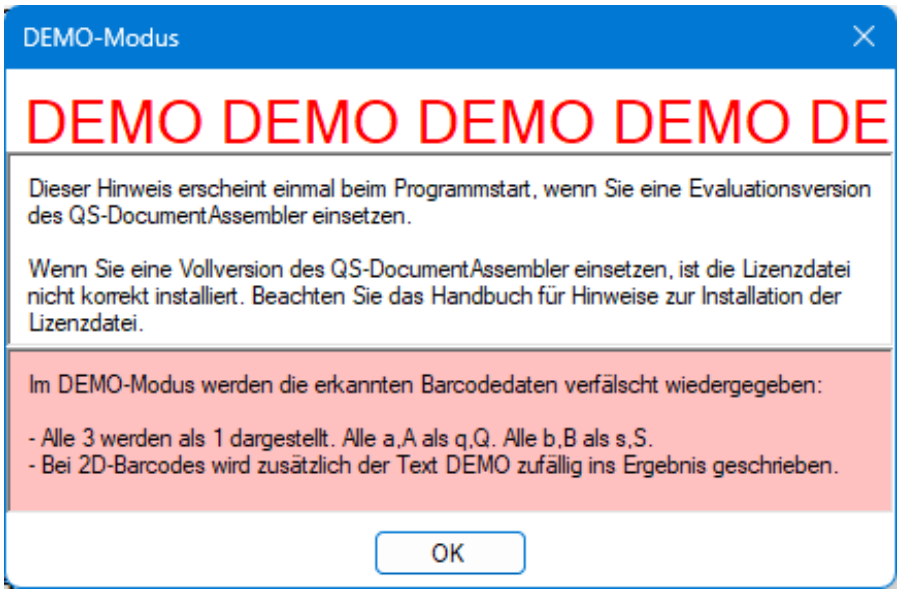
2.3 Demo-Version

Während der Installation wird standardmäßig eine Demo-Lizenz installiert, wenn Sie die Demo-Lizenz bei der Installation nicht durch Ihre gekaufte Lizenz ersetzt haben.

In der Demo-Version können Sie den DocumentAssembler zeitlich unbegrenzt testen. Die einzige Einschränkung in der Benutzung liegt darin, dass die Ergebnisse der Barcodeerkennung systematisch verfälscht werden, d.h. es werden folgende Umsetzungen durchgeführt:

Bei den Ziffern wird die 3 als 1 dargestellt. Alle A,a als Q,q ,alle B,b als S,s. Bei 2D-Barcodes wird außerdem der Text "DEMO" zufällig in das Ergebnis eingeblendet.

Ein Hinweis beim Start des QS-DocumentAssembler macht hierauf noch einmal aufmerksam:



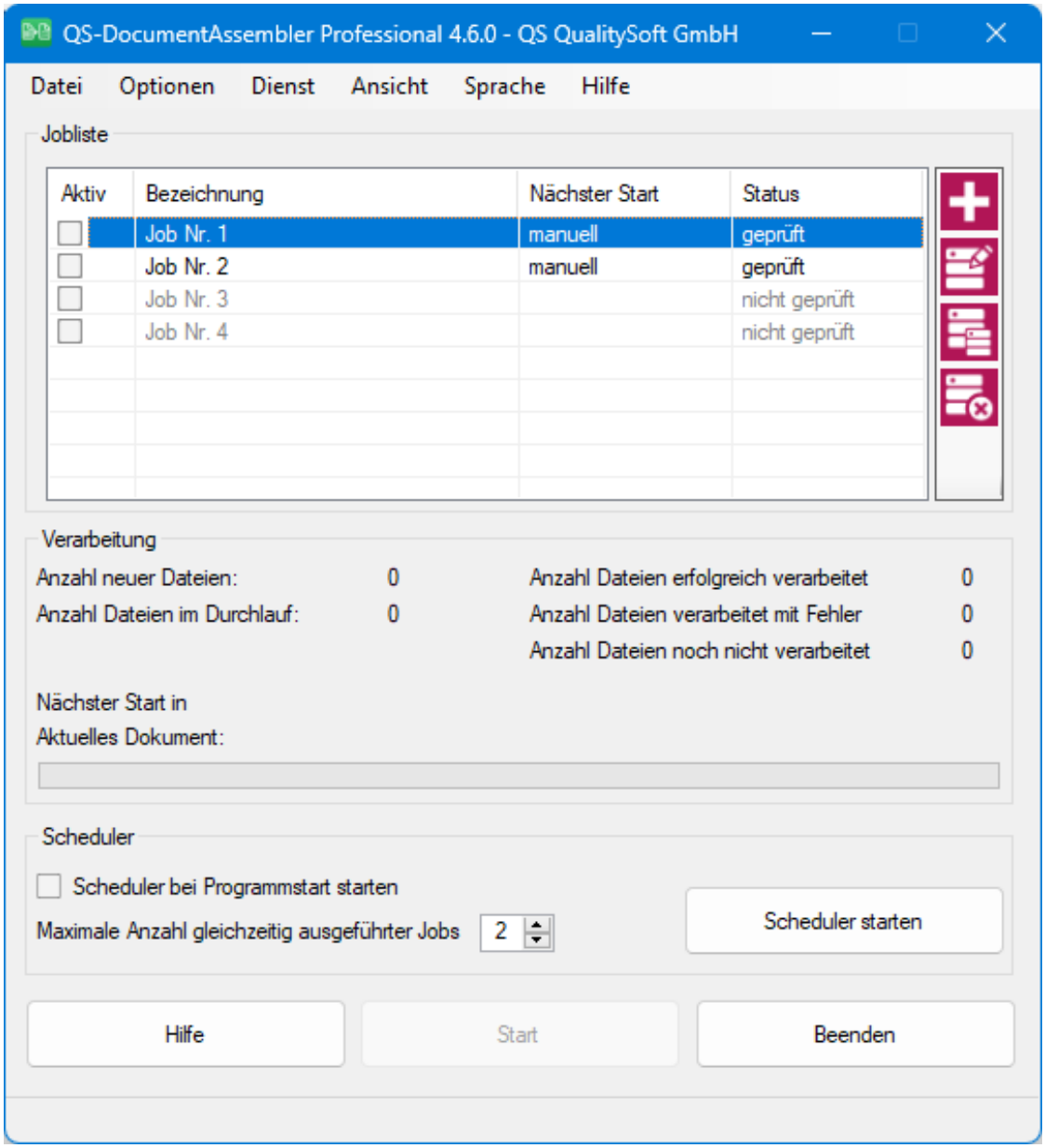
2.4 Vollversion

Wenn Sie eine Vollversion erworben haben, können Sie bei der Installation die Demo-Lizenz durch die Lizenz Ihrer Vollversion im Installationsverzeichnis austauschen (qsbc.lic). Wenn Sie die Vollversion erst nach der Installation der Demo-Version erworben haben, kopieren Sie einfach Ihre Lizenzdatei in das Programmverzeichnis des DocumentAssembler. Sie können dann nach einem Neustart des QS-DocumentAssembler sofort mit der Vollversion weiterarbeiten.

3 Bedienung des QS-DocumentAssembler

3.1 Hauptfenster

Der QS-DocumentAssembler zeigt nach dem Start folgenden Dialog:



3.1.1 Jobliste

Die Basis des DocumentAssembler bildet der Job. Ein Job entspricht einer Konfigurationseinstellung zur Verarbeitung von Dokumenten mit Barcodes.

In der Jobliste werden alle zur Verarbeitung verfügbaren Jobs angezeigt. Damit ein Job ausgeführt werden kann, muss er sich im Status "geprüft" befinden. Dieser Status gibt an, dass der Job alle notwendigen Informationen enthält, um ausgeführt zu werden.

Beim Speichern eines Jobs werden Plausibilitätsprüfungen der eingegebenen Daten durchgeführt (siehe Kapitel 4, Plausibilitätsprüfung). Fehlen noch Daten (z.B. Quellverzeichnis nicht angegeben o.ä.) oder sind die Daten inkonsistent, kann der Job auch ungeprüft gespei-

chert werden. Der Job wird dann im Status "nicht geprüft" angezeigt. Zusätzlich werden diese Jobs ausgegraut dargestellt.

In der Spalte "Nächster Start" können Sie erkennen, wie und wann ein Job ausgeführt wird bzw. ausgeführt werden kann (siehe Kapitel 3.1.2, Ausführung von Jobs).

Rechts neben der Jobliste befindet sich eine vertikale Button-Leiste, über die alle Funktionen zur Bearbeitung von Jobs verfügbar gemacht werden. Sie haben die Möglichkeit einen neuen Job anzulegen, einen Job zu bearbeiten, einen Job zu kopieren sowie einen Job zu löschen.

3.1.2 Ausführung von Jobs

Ein aktiver Job kann immer auf zwei Arten ausgeführt werden. Entweder wird er vom Benutzer manuell über die Schaltfläche <Start> zu einem gewünschten Zeitpunkt gestartet oder er wird automatisch über den Scheduler (siehe Kapitel 3.1.2.2, Scheduler) zu einem definierten Zeitpunkt gestartet.

Die Spalte "Aktiv" zeigt an, ob ein Job ausgeführt werden kann oder nicht. Auch manuell auszuführende Jobs müssen aktiv gesetzt werden, bevor sie ausgeführt werden können. Nicht geprüfte Jobs können nicht aktiv gesetzt werden, da Ihre Einstellungen eine erfolgreiche Ausführung nicht zulassen.

3.1.2.1 Manuelle Ausführung

Über die Schaltfläche <Start> und über das Menü <Ausführung, Start> kann manuell ein Verarbeitungsdurchgang für den ausgewählten Job angestoßen werden, wenn der Job für die manuelle Ausführung vorgesehen ist.

3.1.2.2 Scheduler

Der Scheduler regelt die zeitgesteuerte Ausführung von Jobs, wenn diese für die automatische Verarbeitung vorgesehen sind.

Der Scheduler kann entweder bei Programmstart automatisch oder manuell gestartet werden. Wenn Sie möchten, dass der Scheduler automatisch beim Start des DocumentAssembler gestartet wird, können Sie die Checkbox "Scheduler bei Programmstart starten" aktivieren. Andernfalls müssen Sie den Scheduler über den Button <Scheduler starten> manuell starten.

Nur aktive Jobs werden durch den Scheduler berücksichtigt.

Wenn ein Job für die Ausführung eingeplant ist, wird in der Spalte "Nächster Start" das nächste Ausführungsdatum angezeigt und der Job befindet sich im Status "wartend". Im Bereich "Verarbeitung" wird im Feld "Nächster Start in" die verbleibende Zeit bis zum Start des Jobs angezeigt.

Des Weiteren können Sie hier die maximale Anzahl gleichzeitig ausgeführter Jobs definieren. Dies macht insbesondere dann Sinn, wenn Sie viele Jobs in kurzen Intervallen ausführen. Durch die Begrenzung können sie die maximale Prozessorlast des betreffenden Rechners steuern.

3.1.3 Verarbeitung

Im Bereich "Verarbeitung" wird der Verarbeitungsfortschritt eines laufenden und in der Jobliste ausgewählten Jobs angezeigt. Wenn mehrere Jobs parallel laufen, können Sie durch Auswahl eines anderen laufenden Jobs in der Jobliste den Verarbeitungsfortschritt dieses Jobs anzeigen lassen.

Im Feld "Aktuelles Dokument" wird das z.Z. verarbeitete Dokument inkl. aktuell bearbeiteter Seitenzahl angezeigt. Der Fortschrittsbalken zeigt die Anzahl der verarbeiteten Dokumente im Verhältnis zu ihrer Gesamtanzahl, die im Feld "Anzahl Dateien im Durchlauf" angezeigt wird. Erfolgreich bzw. fehlerhaft verarbeitete Dokumente werden in den Feldern "Anzahl Dateien erfolgreich verarbeitet" bzw. "Anzahl Dateien verarbeitet mit Fehler" angezeigt. Noch nicht verarbeitete Dokumente werden im Feld "Anzahl Dateien noch nicht verarbeitet" angezeigt.

Die Anzeige spiegelt nur den letzten bzw. aktuellen Durchlauf wider. Sie liefert keine historische Sicht auf die Gesamtmenge der verarbeiteten Daten.

Ein Verarbeitungsdurchgang arbeitet immer alle vorhandenen Bilddateien im Quellverzeichnis ab. Dokumente, die während der Verarbeitung dem Quellverzeichnis hinzugefügt werden, werden im Feld "Anzahl neuer Dateien" angezeigt und erst im folgenden Verarbeitungsdurchgang berücksichtigt.

3.1.4 Die Menüleiste

Über die Menüleiste werden z.T. Funktionen bereitgestellt, die sich aus dem Dialog auch direkt aufrufen lassen wie z.B. das Starten und beenden von Jobs aber auch einige erweiterte Funktionen.

Über das Menü <Datei> lässt sich die Anwendung beenden.

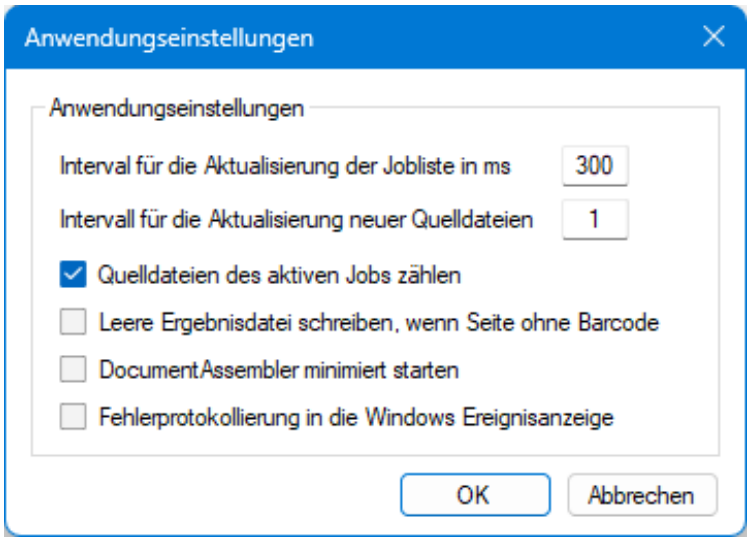
Im Menü <Optionen> können die folgenden Einstellungen vorgenommen werden:

"DocumentAssembler minimiert starten"
: QS-DocumentAssembler startet minimiert. Nach dem Start wird nur das Icon in der Task-Leiste angezeigt.

"Fehlerprotokollierung in das Windows Ereignisprotokoll"
: Standardmäßig protokolliert der QS-DocumentAssembler in das eigene Logfile. Sie können eine zusätzliche Protokollierung in das Windows Ereignisprotokoll einschalten.

"Quelldateien des aktiven Jobs zählen"
: Wenn ein Job auch aus Unterverzeichnissen Dateien verarbeiten soll und sich sehr viele Dateien in den Unterverzeichnissen befinden, kann das Zählen der Dateien zu einer Verlangsamung der Reaktion der Oberfläche kommen. Bei Bedarf können Sie das Zählen der Quelldateien daher ausschalten.

"Anwendungseinstellungen"
: In diesem Dialog können Sie verschiedene Einstellungen die Anwendung betreffend ändern.



Intervall für die Aktualisierung der Jobliste in ms

Die Anzeige der Daten eines Jobs wird zeitgesteuert alle 300 ms aktualisiert. Bei Bedarf können Sie den Wert hier hochsetzen.

Intervall für die Aktualisierung der Quelldateien

Die Anzahl der Quelldateien wird in jedem zeitgesteuerten Durchlauf aktualisiert. Wenn die Quelldateien nur in jedem dritten Durchlauf aktualisiert werden sollen, setzen Sie den Wert auf 3.

Quelldateien des aktiven Jobs zählen

S.o.

Leere Ergebnisdatei schreiben, wenn Seite ohne Barcode

Normalerweise wird nur dann eine Ergebnisdatei geschrieben, wenn ein Barcode gefunden wurde. Wenn Sie diese Option auswählen wird eine Ergebnisdatei mit einer Größe von 0 Byte geschrieben, wenn kein Barcode gefunden wurde.

DocumentAssembler minimiert starten

S.o.

Fehlerprotokollierung in das Windows Ereignisprotokoll

S.o.

Über das Menü <Dienst> lässt sich der Dienst installieren und deinstallieren. Außerdem können Sie hier den Dienst starten und einen laufenden Dienst stoppen. Um diese Funktionen ausführen zu können, muss der DocumentAssembler mit Administratorrechten gestartet werden. Falls diese Funktion ohne Administratorrechte ausgeführt wird, werden Sie gefragt, ob der DocumentAssembler mit Administratorrechten neu gestartet werden soll.

Das Menü <Ansicht> dient zur Anzeige verschiedener Protokolldateien. Ein Suchen nach diesen Dateien in der Verzeichnisstruktur entfällt somit.

"Protokolldatei"
Anzeige der zum ausgewählten Job zugehörigen Protokolldatei.

"Ergebnisdatei"
Anzeige der zum ausgewählten Job zugehörigen Ergebnisdatei.

"Fehlerdatei"
Anzeige der zum ausgewählten Job zugehörigen Fehlerdatei.

"Allgemeine Fehlerdatei"
Anzeige der zur Anwendung zugehörigen Fehlerdatei.

Im Menü <Sprache> können Sie zwischen den verfügbaren Sprachen umschalten. Z.Z. werden Deutsch und Englisch unterstützt.

Im Menü <Hilfe> können Sie das Handbuch des QS-DocumentAssembler als PDF-Dokument öffnen sowie ein PDF-Dokument mit allgemeinen Informationen zur Erkennung von Barcodes. Außerdem lässt sich hier über das Menü <Fernwartung> der "Quick Support" des Teamviewers zur Supportunterstützung starten.

Außerdem lässt sich hier ein "Info-Dialog" starten, der Hinweise auf Programmversionen und Herstellerkontaktdaten beinhaltet.

3.1.5 Die Erstellung eines neuen Jobs

Ein neuer Job kann zu jeder Zeit über den Button "Neuer Job" angelegt werden und anschließend im Dialog "Einstellungen" bearbeitet werden. Die Bearbeitung erfolgt über mehrere auf dem Dialog befindliche Registerkarten.

3.1.6 Die Bearbeitung von Jobs

Über die Button-Leiste rechts neben der Jobliste lassen sich neue Jobs anlegen. Außerdem können bestehende Jobs bearbeitet, kopiert oder gelöscht werden.

Neuer Job: Ein neuer Job wird erstellt. Der Job wird mit Standard-Parametern vorbelegt.

Job bearbeiten: Ruft den Dialog für die Bearbeitung des markierten Jobs auf.

Job kopieren: Duplizieren des aktuell markierten Jobs.

Job löschen: Löschen des aktuell markierten Jobs.

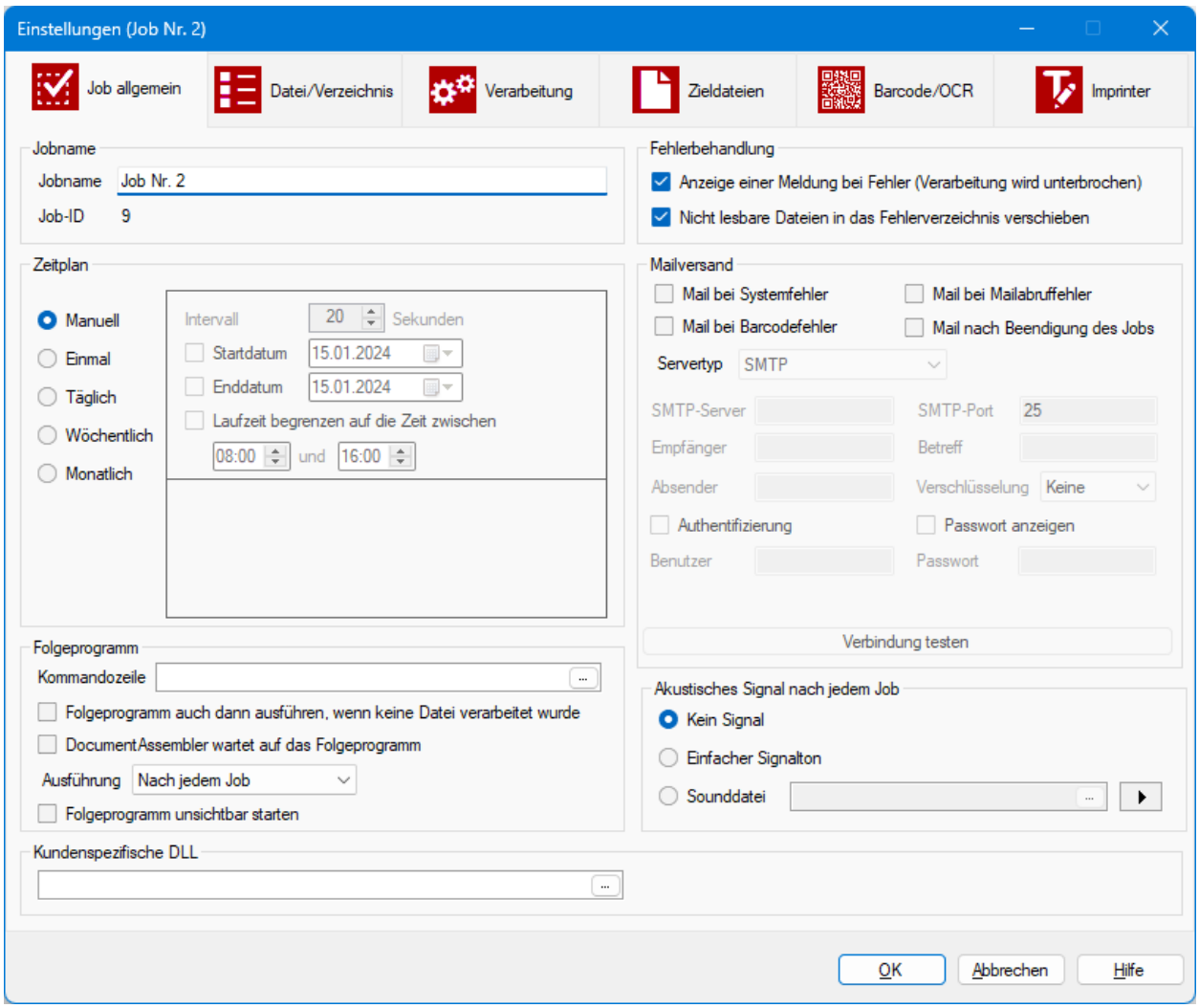
In Abhängigkeit des Status eines ausgewählten Jobs sind eventuell nicht alle Funktionen verfügbar. So kann z.B. aktiv geschalteter Job nicht bearbeitet oder gelöscht werden.

Die Bearbeitung bestehender Jobs lässt sich auch über einen Doppelklick auf den Job auslösen. Der Job darf dabei nicht aktiv sein.

3.2 Einstellungen

Über den Button "**Job bearbeiten**" kommen Sie in die Einstellungen des Jobs.

3.2.1 Registerkarte "Job allgemein"



Jobname

Jeder Job muss einen eindeutigen Namen bekommen. Doppelte Namen sind aus technischen Gründen nicht möglich. Unter dem Namen des Jobs wird die interne ID des Jobs angezeigt.

Zeitplan

Der DocumentAssembler bietet die Möglichkeit, die Jobs manuell sowie zeitgesteuert in verschiedenen Zeitfenstern auszuführen.

Der Scheduler überwacht die Liste der abzuarbeitenden Jobs und startet diese zum definierten Zeitpunkt. Es wird nur die maximale Anzahl gleichzeitig ausgeführter Jobs gestartet. Weitere zu startende Jobs werden in eine Warteschlange gestellt und dann gestartet, wenn die Anzahl der parallel abzuarbeitenden Jobs es zulässt.

Wenn es keine weiteren Ausführungszeitpunkte für einen Job mehr gibt, wird ein Job deaktiviert und bekommt den Status "abgelaufen".

Wenn der Scheduler zu einem Zeitpunkt gestartet wird, bei dem ein Job bereits abgelaufen ist, wird der Job nicht mehr ausgeführt, sondern sofort mit dem Status "abgelaufen" versehen.

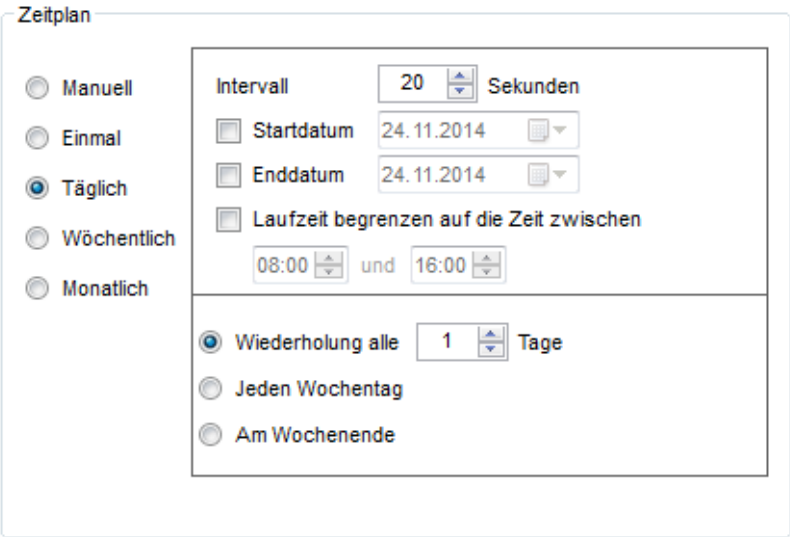
Im Folgenden werden die einzelnen zeitgesteuerten Auslösezeitpunkte und ihre Einstellungen beschrieben.

- **Manuell**
  Hier sind keine Einstellungen möglich, da der Job manuell aus der Jobliste zu starten ist und nicht über den Scheduler ausgeführt werden kann.

- **Einmal**

  Wenn Sie die Optionsschaltfläche **Einmal** aktivieren, müssen Sie ein Startdatum und eine Startzeit auswählen, zu dem der Job gestartet werden soll. Das Datum muss in der Zukunft liegen. Nach erfolgreicher Ausführung bzw. nach Ablauf der Zeit wird der Job deaktiviert und befindet sich im Status "abgelaufen".

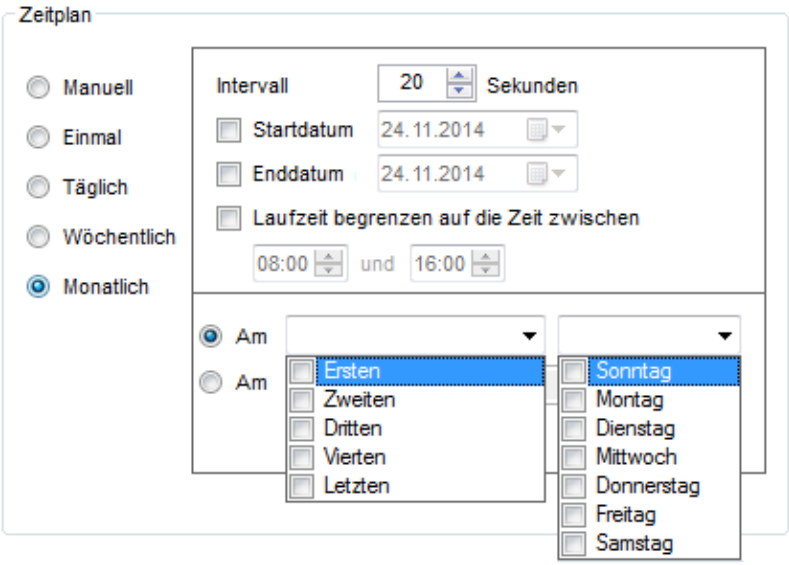
- **Täglich**



  Bei der täglichen Ausführung eines Jobs müssen Sie das Ausführungsintervall in Sekunden angeben. Das Ausführungsintervall bestimmt, in welchen Abständen ein Job

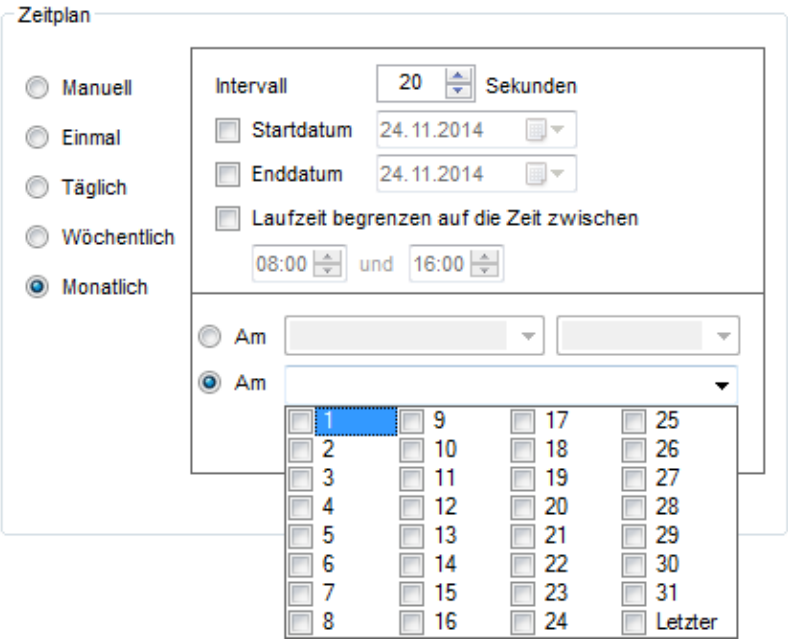
vom Scheduler gestartet wird.
Außerdem können Sie ein Start- und Enddatum festlegen, müssen es jedoch nicht. Wenn Sie kein Start- oder Enddatum festlegen, wird der Job ab sofort bzw. dauerhaft eingeplant.
Die Ausführung kann tageweise (jeder Tag, jeder zweite Tag, etc.) oder auf einen Wochentag bzw. auf das Wochenende beschränkt werden. Die Laufzeit kann zusätzlich über den Tag begrenzt werden.

- **Wöchentlich**

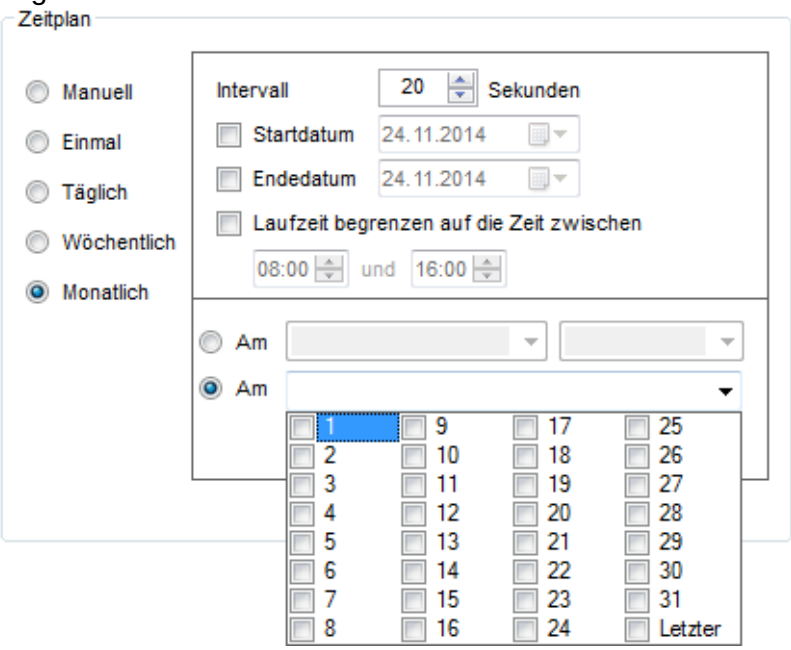


Bei der wöchentlichen Ausführung eines Jobs müssen Sie das Ausführungsintervall in Sekunden angeben. Das Ausführungsintervall bestimmt, in welchen Abständen ein Job vom Scheduler gestartet wird.
Außerdem können Sie ein Start- und Enddatum festlegen, müssen es jedoch nicht. Wenn Sie kein Start- oder Enddatum festlegen, wird der Job ab sofort bzw. dauerhaft eingeplant.
Die Ausführung kann wochenweise (jede Woche, jede zweite Woche, etc.) und auf bestimme Wochentage beschränkt werden. Die Laufzeit kann zusätzlich über den Tag begrenzt werden.

- **Monatlich**



Bei der monatlichen Ausführung eines Jobs müssen Sie das Ausführungsintervall in Sekunden angeben. Das Ausführungsintervall bestimmt, in welchen Abständen ein Job vom Scheduler gestartet wird.
Außerdem können Sie ein Start- und Enddatum festlegen, müssen es jedoch nicht. Wenn Sie kein Start- oder Enddatum festlegen, wird der Job ab sofort bzw. dauerhaft eingeplant.
Die Ausführung kann entweder wochenweise und innerhalb der Woche auf einzelne Tage oder



auf einzelne Tage innerhalb des Monats beschränkt werden.
Die Laufzeit kann zusätzlich über den Tag begrenzt werden.

Folgeprogramm

Der QS-DocumentAssembler kann am Ende eines Jobs (wenn alle Dateien aus dem Quellverzeichnis bearbeitet sind) oder nach der Verarbeitung jeder einzelnen Datei ein Folgeprogramm starten.

"Kommandozeile"
: Der gewünschte Programmaufruf wird hier eingetragen.

"Folgeprogramm auch dann ausführen, wenn keine Datei verarbeitet wurde"
: Gelegentlich ist es gewünscht, das Folgeprogramm auch dann auszuführen, wenn in einem Durchlauf keine Dateien verarbeitet wurden.

"DocumentAssembler wartet auf das Folgeprogramm"
: Bestimmt, ob der QS-DocumentAssembler auf das Ende der Folgeverarbeitung warten soll oder nicht.

"Ausführung"
: Das Folgeprogramm kann nach jedem Job oder nach jedem Dokument ausgeführt werden.

"Folgeprogramm unsichtbar starten"
: Das Folgeprogramm kann sichtbar gestartet werden, wenn ggf. noch eine Benutzerinteraktion stattfinden soll oder unsichtbar, so dass das Fenster des ausführenden Programmes nicht angezeigt wird. Dies verhindert ggf. auch eine störende Unterbrechung, wenn an dem Rechner parallel gearbeitet werden soll.

Kundenspezifische DLL

Wenn Sie spezielle Funktionen benötigen, die der DocumentAssembler nicht zur Verfügung stellt, sprechen Sie uns an. Wir erstellen Ihnen gern ein Angebot für die Implementierung kundenspezifischer Funktionen.

Fehlerbehandlung

Legen Sie das Verhalten des QS-DocumentAssembler beim Auftreten von Fehlern fest.

"Anzeige einer Meldung bei Fehler (Verarbeitung wird unterbrochen)"
: Wenn bei der Verarbeitung der Dateien Fehler auftreten, kann eine entsprechende Meldung angezeigt werden. Die Verarbeitung wird damit unterbrochen. Diese Option ist standardmäßig ausgeschaltet.

"Nicht lesbare Dateien in das Fehlerverzeichnis verschieben"
: Wenn Dateien nicht gelesen werden können, kann man sie in das Fehlerverzeichnis verschieben lassen. Diese Option ist standardmäßig eingeschaltet. Wenn diese Option nicht aktiviert ist, wird die gesamte Verarbeitung abgebrochen. Die Verarbeitung kann erst dann fortgesetzt werden, wenn diese Datei entfernt wurde.

Mailversand

Sie haben die Möglichkeit, sich beim Auftreten eines Problems durch den QS-DocumentAssembler eine E-Mail senden zu lassen. Somit werden sie umgehend über technische Probleme in der Verarbeitung informiert. Sie haben die Möglichkeit, mehrere Empfänger durch ein Semikolon separiert einzutragen. Unterstützt werden SMTP-Server sowie der Versand von Mails über Outlook365.

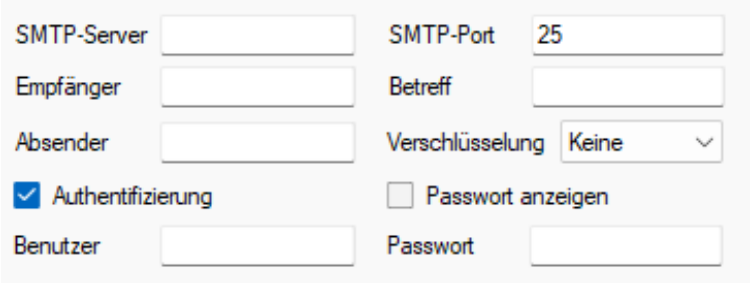
" Mail bei Systemfehler"
Senden Sie eine Mail, wenn ein Programmfehler aufgetreten ist.

" Mail bei Barcodefehler"
Senden Sie eine Mail, wenn ein Barcode auf Basis der Einstellungen nicht gelesen werden konnte oder ein Dokument in das Fehlerverzeichnis verschoben wird.

"Mail nach Beendigung des Jobs"
Senden Sie eine Mail, wenn der Job alle Dateien des Quellverzeichnisses verarbeitet hat.

" Mail bei Mailabruffehler"
Senden Sie eine Mail, wenn Sie Dokumente aus dem Postkorb abholen und ein Fehler beim Abrufen von bzw. beim Speichern von Anhängen aufgetreten ist.

Servertyp SMTP



"SMTP-Server"
SMTP-Server, über den die Mail versandt werden soll.

"SMTP-Port"
Der Port, über den Sie mit Ihrem SMTP-Server kommunizieren.

"Empfänger"
Mailadresse des Empfängers der Benachrichtigung. Es können mehrere Empfänger durch ein Semikolon getrennt eingegeben werden.

"Betreff"
Falls Sie auf mehreren Computern den QS-DocumentAssembler einsetzen, können Sie über die Betreff-Zeile eine Differenzierung des Absenders vornehmen.

"*Absender*"
Mailadresse des Absenders der Benachrichtigung.

"*Verschlüsselung*"
Verschlüsselung, mit der der SMTP-Server Mails verschlüsselt.

"*Authentifizierung*"
Wenn Ihr SMTP-Server eine anonyme Anmeldung unterstützt, können Sie diese Option deaktivieren.

"*Benutzer*"
Benutzername für die Anmeldung am SMTP-Server.

"*Passwort*"
Passwort für die Anmeldung am SMTP-Server.

Servertyp Outlook356

"Mandanten-ID"
Hierbei handelt es sich um eine spezifische Instanz im Azure Active Directory (AAD), in der Benutzerkonten und Gruppen gespeichert werden. Bei der ID handelt es sich um eine eindeutige GUID, welche auch als Tenant bezeichnet wird.

"Anwendungs-ID"
Manuell im AAD erstellte Anwendung (GUID) für den DocumentAssembler, um über Outlook356 auf einen Postkorb zugreifen zu können.

"Absender"
Zeigt den Absender an, nachdem ein Outlook365-Benutzer für den Versand von Mails konfiguriert ist. Dieses Feld ist schreibgeschützt.

"Betreff"
Falls Sie auf mehreren Computern den QS-DocumentAssembler einsetzen, können Sie über die Betreff-Zeile eine Differenzierung des Absenders vornehmen.

"Empfänger"
Mailadresse des Empfängers der Benachrichtigung. Es können mehrere Empfänger durch ein Semikolon getrennt eingegeben werden.

"Outlook-Postfach ist aktiviert."
Zeigt an, ob bereits ein Postkorb für den Versand von Mails konfiguriert wurde. Dieses Feld ist schreibgeschützt.

Die Konfiguration von Outlook365 für das Senden und Abrufen von Mails wird in Kapitel 3.2.2.1 näher beschrieben.

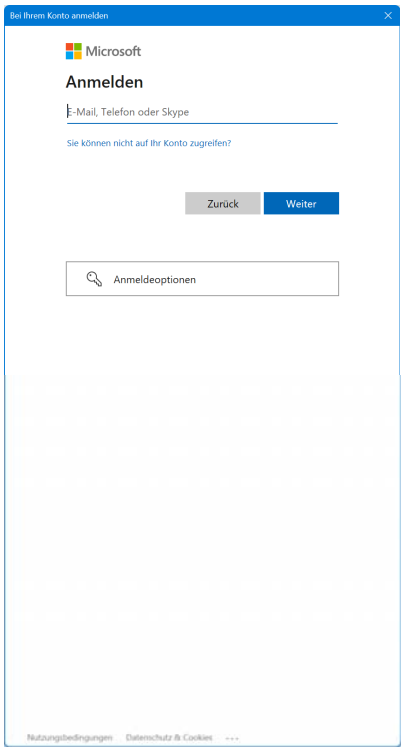
Neuen Benutzer aktivieren

Um für den DocumentAssembler einen Outlook365-Benutzer einzurichten, benötigen Sie Ihre Mandanten-ID und eine Anwendungs-ID. Wo Sie die Mandanten-ID finden und wie Sie eine Anwendungs-ID erstellen erfahren Sie im Kapitel 3.2.2.1

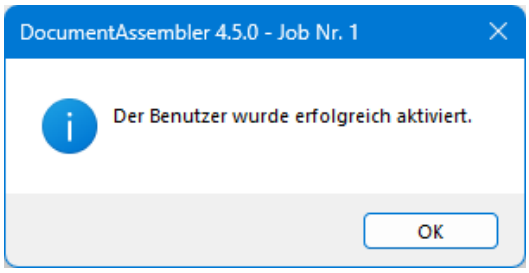
Wählen Sie als Servertyp "Outlook365" und geben Sie Mandanten-ID und Anwendungs-ID ein und klicken auf den Button "Neuen Benutzer aktivieren".

Dann erscheint der Anmeldedialog für Outlook365. Dort müssen Sie sich einmalig mit dem gewünschten Benutzer und Passwort bei Outlook 365 anmelden.

Falls Sie auf dem Rechner schon einmal in Outlook365 angemeldet waren kann es auch sein, dass Sie nur den Benutzer auswählen müssen.



Bei erfolgreicher Anmeldung erhalten Sie die Meldung:

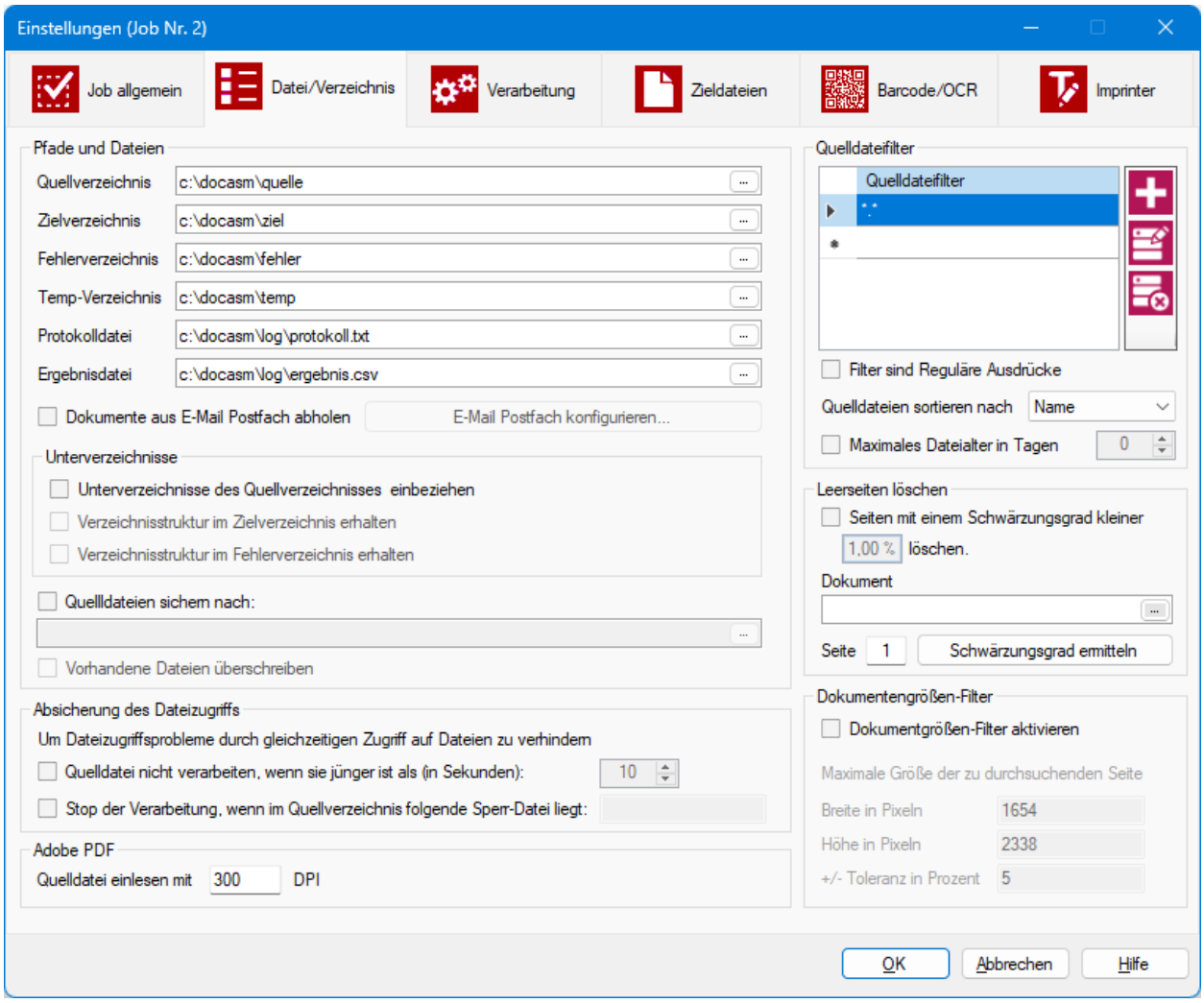


Während der Anmeldung wird von Outlook365 ein Token generiert, welches zukünftig für die Anmeldung benutzt wird. Dieses Token wird mit AES-256 verschlüsselt in der Datenbank gespeichert. Vor jedem Zugriff auf den Postkorb wird das Token aktualisiert.

Akustisches Signal nach jedem Job

Sie haben die Möglichkeit, nach jedem Durchlauf des Jobs ein Signal auszugeben. Sie können hierbei wählen, ob "kein Signal" ausgegeben werden soll, ein "einfacher Signalton" oder eine beliebige Sounddatei. Mit der Ausführen-Schaltfläche können Sie prüfen, ob geeignete Treiber und Ausgabegeräte im System installiert sind.

3.2.2 Registerkarte "Datei/Verzeichnis"



Pfade und Dateinamen

"*Quellverzeichnis*"

Das *Quellverzeichnis* bezeichnet das Verzeichnis, in dem die zu verarbeitenden Dokumente gesucht werden.

"*Zielverzeichnis*"

Das Zielverzeichnis verweist auf das Verzeichnis, in dem Dokumente nach der Verarbeitung abgelegt werden.

"*Fehlerverzeichnis*"

Wenn Sie die Verarbeitung so konfigurieren, dass Seiten ohne Barcode aussortiert werden, werden die Seiten in dieses Verzeichnis verschoben.

"*Protokolldatei* "

In der Protokolldatei wird die Verarbeitung jeder einzelnen Seite dokumentiert. Sie ist hier komplett mit Pfadangabe anzugeben. Für den Dateinamen des Protokolls besteht die Möglichkeit, Datumsparameter (Tag, Monat, Jahr) in den Dateinamen aufzunehmen. Folgende Parameter sind gültig:

%YYYY%	Jahr vierstellig. z.B. 2014
%YY%	Jahr zweistellig, z.B. 14
%MMMM%	Monat Langform, z.B. "September"
%MMM%	Monat als Kürzel, z.B. "Sep"
%MM%	Monat als Ziffer mit führender Null, z.B. "07"
%M%	Monat einstellig
%DDDD%	Tag Langform, z.B. "Dienstag"
%DDD%	Tag als Kürzel, z.B. "Di"
%DD%	Tag als Ziffer mit führender Null, z.B. "01"
%D%	Tag, einstellig
%TDJ%	Tag des Jahres

Der Eintrag könnte z.B. lauten:

c:\docasm\log\Protokoll_%MM%%YYYY%.log
Mit dieser Einstellung würde z.B. monatlich eine neue Logdatei erstellt werden.

"*Ergebnisdatei*"

Für den Fall, dass alle Barcode-Informationen in einer Datei gesammelt werden sollen, kann hier der Dateiname spezifiziert werden. Für den Dateinamen der Ergebnisdatei können ebenfalls die o.g. Datumsparameter verwendet werden.

"*Temp. Verzeichnis*"

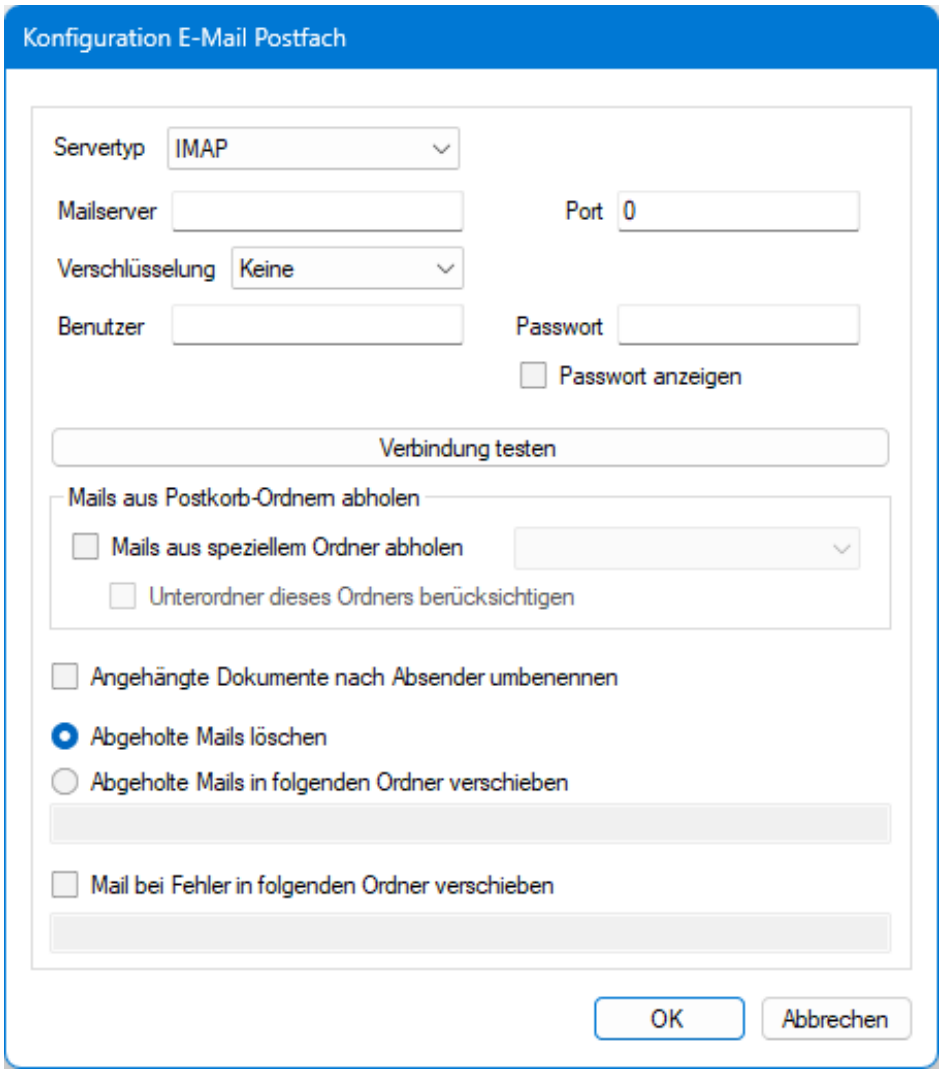
Im Rahmen der Verarbeitung erstellte temporäre Dateien werden in diesem Verzeichnis gespeichert.

Zur Bildung aller Verzeichnisnamen können auch eine oder auch mehrere Umgebungsvariablen benutzt werden. Die Umgebungsvariablen sind mit einem Prozentzeichen (%) beidseitig zu maskieren. Die Pfadangabe für das Quellverzeichnis könnte z.B. lauten:

%SystemDrive%\docasm\quelle\

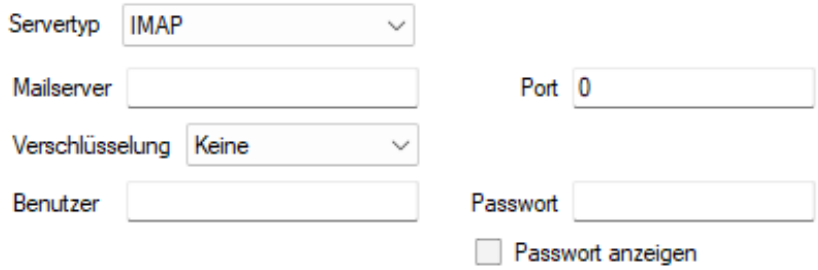
Dokumente aus einem E-Mail-Postfach abholen

Neben dem Lesen von Quelldokumenten aus einer Verzeichnisstruktur unterstützt der DocumentAssembler auch das Abholen von Dokumenten (als Anhang) aus E-Mail-Postfächern.



"Servertyp"
Es werden IMAP, POP3, Exchange und Outlook365 unterstützt.

Servertyp IMAP/POP3/EXCHANGE



"Mailserver"
Name des Servers, über den die Mails abgerufen werden.

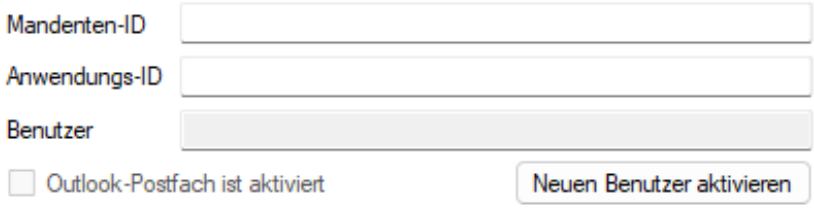
" Port"
Der Port, über den Sie mit Ihrem Mail-Server kommunizieren.

"Verschlüsselung"
Es werden SSL/TLS, STARTTLS und keine Verschlüsselung unterstützt.

"*Benutzer*"
Benutzername für die Anmeldung am Mail-Server.

"*Passwort*"
Passwort für die Anmeldung am Mail-Server.

Servertyp Outlook365



"Mandanten-ID" Hierbei handelt es sich um eine spezifische Instanz im Azure Active Directory (AAD), in der Benutzerkonten und Gruppen gespeichert werden. Bei der ID handelt es sich um eine eindeutige GUID, welche auch als Tenant bezeichnet wird.

"Anwendungs-ID"
Manuell im AAD erstellte Anwendung (GUID) für den DocumentAssembler, um über Outlook356 auf einen Postkorb zugreifen zu können.

"Benutzer"
Benutzername für die Anmeldung am Postkorb. Wird erst angezeigt, nachdem ein Outlook365-Benutzer für den Abruf von Mails konfiguriert ist. Dieses Feld ist schreibgeschützt.

"Outlook-Postfach ist aktiviert."

Zeigt an, ob bereits ein Postkorb für den Abruf von Mails konfiguriert wurde. Dieses Feld ist schreibgeschützt.

Die Konfiguration von Outlook365 für das Senden und Abrufen von Mails wird in Kapitel 3.2.2.1 Die Konfiguration von Outlook365 zum Senden und Abrufen von Mails näher beschrieben.

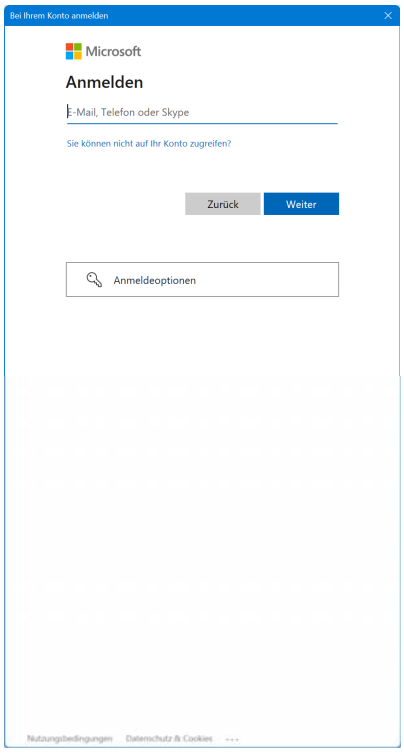
Neuen Benutzer aktivieren

Um für den DocumentAssembler einen Outlook365-Benutzer einzurichten, benötigen Sie Ihre Mandanten-ID und eine Anwendungs-ID. Wo Sie die Mandanten-ID finden und wie Sie eine Anwendungs-ID erstellen erfahren Sie im Kapitel 3.2.2.1 Die Konfiguration von Outlook365 zum Senden und Abrufen von Mails.

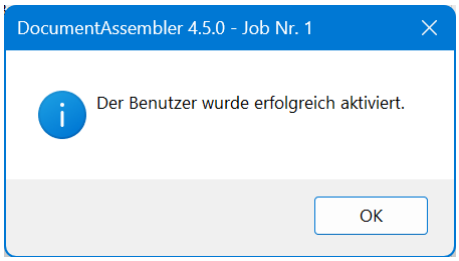
Wählen Sie als Servertyp "Outlook365" und geben Sie Mandanten-ID und Anwendungs-ID ein und klicken auf den Button "Neuen Benutzer aktivieren".

Dann erscheint der Anmeldedialog für Outlook365. Dort müssen Sie sich einmalig mit dem gewünschten Benutzer und Passwort bei Outlook 365 anmelden.

Falls Sie auf dem Rechner schon einmal in Outlook365 angemeldet waren kann es auch sein, dass Sie nur den Benutzer auswählen müssen.



Bei erfolgreicher Anmeldung erhalten Sie die Meldung:



Während der Anmeldung wird von Outlook365 ein Token generiert, welches zukünftig für die Anmeldung benutzt wird. Dieses Token wird mit AES-256 verschlüsselt in der Datenbank gespeichert. Vor jedem Zugriff auf den Postkorb wird das Token aktualisiert.

Mails aus Postkorb-Ordnern abholen

Wenn nichts anderes definiert ist, werden die Mails standardmäßig aus dem Eingangsordner abgeholt. Es besteht allerdings die Möglichkeit, die Mails auch aus speziellen Ordnern des Postkorbes abzuholen.

Beim erstmaligen Öffnen der Combobox für die Postkorb-Ordner wird der Standard-Postkorb ausgewählt.

"Mails aus speziellem Ordner abholen"
: Geben Sie hier einen beliebigen Ordner des Postkorbes an, aus dem die Mails abgeholt werden sollen.

"Unterordner dieses Ordners berücksichtigen"
: Hier können sie angeben, ob vorhandene Unterordner des vorher definierten Ordners beim Abholen der Mails berücksichtigt werden sollen.
Bitte beachten Sie: Wenn sie diese Option ausgewählt haben, werden die angehängten Dateien im definierten Quellverzeichnis analog in Unterordner entsprechend der Namen der der speziellen Ordner gespeichert. Um diese Dateien anschließend mit dem DocumentAssembler zu verarbeiten muss die Option "Unterverzeichnisse des Quellverzeichnisses einbeziehen" ausgewählt sein.

"*Abgeholte Mails löschen*"
: Nach dem Abholen der Mails und dem Speichern der Anhänge wird die jeweilige Mail gelöscht.

"*Abgeholte Mails in folgenden Ordner verschieben*"
: Bei Bedarf kann eine abgeholte Mail nach dem Speichern der Anhänge in einen definierten Ordner verschoben werden. Das Verschieben der Mails in einen Ordner wird nur von IMAP-Servern unterstützt. Falls der Ordner auf dem Server nicht existiert, wird er durch den DocumentAssembler angelegt. Ggf. müssen sie den angelegten Ordner in Ihrem Mailprogramm abonnieren, um ihn zu sehen.

"*Mail bei Fehler in folgenden Ordner verschieben*"
: Wenn es während des Abholens einer Mail oder beim Speichern der angehängten Dateien zu einem Fehler kommt, kann die Mail samt Anhang in den angegebenen Ordner verschoben werden. Falls der Ordner nicht existiert wird versucht, ihn vorher anzulegen. Ggf. müssen sie den angelegten Ordner in Ihrem Mailprogramm abonnieren, um ihn zu sehen.

Wenn Dateianhänge nicht dem Dateifilter entsprechen, werden diese beim Speichern ignoriert.

E-Mails ohne Anhang werden in der Protokolldatei protokolliert und anschließend je nach Einstellung gelöscht oder verschoben.

Da POP3-Server keine Ordner unterstützen, gelten die Ordner-Optionen nur für IMAP/Exchange-Server.

Unterverzeichnisse

"*Unterverzeichnisse des Quellverzeichnisses einbeziehen*"
: Wenn sich zu verarbeitende Dokumente auch in Unterverzeichnissen des Quellverzeichnisses befinden, können diese auch in die Verarbeitung eingeschlossen werden.

"*Verzeichnisstruktur im Zielverzeichnis erhalten*"
: Die verarbeiteten Dokumente werden normalerweise direkt in das Zielverzeichnis geschrieben. Wenn auch Dateien aus Unterverzeichnissen des Quellverzeichnisses verarbeitet werden, können Sie die Verzeichnisstruktur der jeweiligen Unterverzeichnisse im Zielverzeichnis erhalten.

"*Verzeichnisstruktur im Zielverzeichnis erhalten*"
: Die fehlerhaften Dokumente werden normalerweise direkt in das Fehlerverzeichnis geschrieben. Wenn auch Dateien aus Unterverzeichnissen des Quellverzeichnisses verarbeitet werden, können Sie die Verzeichnisstruktur der jeweiligen Unterverzeichnisse im Fehlerverzeichnis erhalten.

Backup der Quelldateien

"*Quelldateien sichern nach*"
: Quelldateien werden im Standard nach der Verarbeitung gelöscht. Sie haben die Möglichkeit, die Quelldateien vor der Verarbeitung in das angegebene Verzeichnis zu kopieren.

"*Vorhandene Dateien überschreiben*"
: Beim Backup der Originaldateien kann angegeben werden, ob bestehende Dateien überschrieben werden sollen. Ansonsten wird eine Fehlermeldung produziert, wenn versucht wird, eine bestehende Datei zu überschreiben.

Absicherung des Dateizugriffs

Beim Lesen, Bewegen und Umbenennen der Quelldateien kann es zu Zugriffskonflikten mit anderen Programmen kommen.

Häufigster Fall: Die Datei ist von der erzeugenden Scanner-Software bereits angelegt, aber noch nicht komplett geschrieben, während der QS-DocumentAssembler bereits auf die Datei zugreift Um solche Konflikte zu vermeiden, bietet der QS-DocumentAssembler zwei Mechanismen:

"Quelldatei nicht verarbeiten, wenn sie jünger ist als (in Sekunden)"
: Es wird anhand des Datums des letzten Zugriffs geprüft, ob die Datei ein Mindestalter erreicht hat. Ist dies nicht der Fall, wird die Verarbeitung gestoppt.
Hinweis: Wenn die Dateien auf einem anderen Rechner erzeugt werden als auf dem QS-DocumentAssembler Rechner, dürfen die Rechner-Uhrzeiten nicht unterschiedlich sein.

"Stopp der Verarbeitung, wenn im Quellverzeichnis folgende Sperr-Datei liegt"

Wenn die andere in Konflikt stehende Anwendung die Möglichkeit bietet, eine Sperrdatei zu schreiben, kann der QS-DocumentAssembler auf das Vorhandensein einer solchen prüfen und die Verarbeitung beenden. Bitte beachten Sie, dass Sie hier lediglich den Dateinamen ohne den zugehörigen Dateipfad eingeben müssen.

Adobe PDF

"Quelldatei einlesen mit n DPI"

Zur Barcodeerkennung werden die Seiten der PDF-Dokumente intern in Bilddateien gewandelt und dabei wird über den angegebenen Wert die Auflösung bestimmt. Zur sicheren Barcodeerkennung sollten Sie einen Wert von 200 DPI nur im Sonderfall unterschreiten. Bei Werten deutlich über 200 DPI müssen Sie zur guten Erkennung eventuell Barcode-Parameter, vor allem die "Ruhezone", anpassen.

Quelldateifilter

Wenn Sie den Zugriff auf bestimmte Dateitypen eingrenzen wollen, können Sie hier einen Dateifilter in der Form "*.*", "*.tif" oder "*.pdf" definieren. Sie können hier beliebig viele Dateifilter definieren.

"Filter sind reguläre Ausdrücke"

Alternativ können Sie komplexere Dateifilter über reguläre Ausdrücke definieren. Bei regulären Ausdrücken sind sie z.B. nicht auf die Dateiendung beschränkt und können mehr als einen Filter definieren.

"Quelldateien sortieren nach"

Quelldateien können nach Namen oder nach dem Alter (Erstellungsdatum der Datei) sortiert werden. *"Maximales Dateialter in Tagen"*
Schränken sie die Quelldateiauswahl auf Dateien ein, die ein bestimmtes Erstellungsdatum nicht unterschreiten.

Für die Verarbeitung von TIF- und PDF-Dokumenten gleichzeitig können Sie z.B. folgenden Regulären Ausdruck verwenden: ^.*\.tif$|^.*\.pdf$

Leerseiten löschen:

Es besteht die Möglichkeit, Leerseiten vor der Barcodeerkennung aus den Dokumenten zu löschen. Dies ist zum Beispiel praktisch, wenn Vorder- und Rückseiten gescannt werden, die leeren Rückseiten aber nicht benötigt werden. Die Entscheidung darüber, ob eine Seite leer ist, wird anhand des Schwärzungsgrades einer Seite getroffen.

"Seiten mit einem Schwärzungsgrad kleiner n% löschen"

Geben Sie hier den Schwärzungsgrad in Prozent an, bis zu dem eine Seite bei Ihnen als Leerseite gilt. Sie können den Schwärzungsgrad auf zwei Nachkommastellen angeben

"Schwärzungsgrad ermitteln"

Hier können Sie für Ihre zu löschenden Leerseiten deren Schwärzungsgrad in Prozent ermitteln. Wählen Sie die gewünschte Datei (Drag and Drop wird unterstützt) und Seite aus und klicken auf den Button <Schwärzungsgrad ermitteln>. Diesen Wert können Sie dann in das o.g. Feld übertragen.

Dokumentengrößen-Filter

Bei einigen Kunden des QS-DocumentAssembler hat sich diese Funktion als nützlicher Zusatz herausgestellt: Wenn Dokumente unterschiedlichen Formats verarbeitet werden, z.B. auch große technische Zeichnungen, so kann die Verarbeitung beschleunigt werden, wenn auf diesen großen Bilddateien keine Barcodes gesucht werden, sondern z.B. nur auf DIN A4 großen Trennblättern. Die Seitengröße kann eingestellt (Angaben in Pixel) und eine Toleranz definiert werden, weil einige Scanner mit "automatischer Papiergrößenerkennung" Bilddateien leicht variierender Größe liefern.

Alle Dokumente, die außerhalb des angegebenen Bereichs liegen (also größer oder kleiner sind), werden behandelt als wäre kein Barcode gefunden worden.

3.2.2.1 Die Konfiguration von Outlook365 zum Senden und Abrufen von Mails

Für die Konfiguration von Outlook365 zum Senden und Abrufen vom Mails benötigen Sie Ihre Mandanten-ID, eine Anwendungs-ID sowie einen Benutzer mit Kennwort zur Anmeldung an Outlook365.

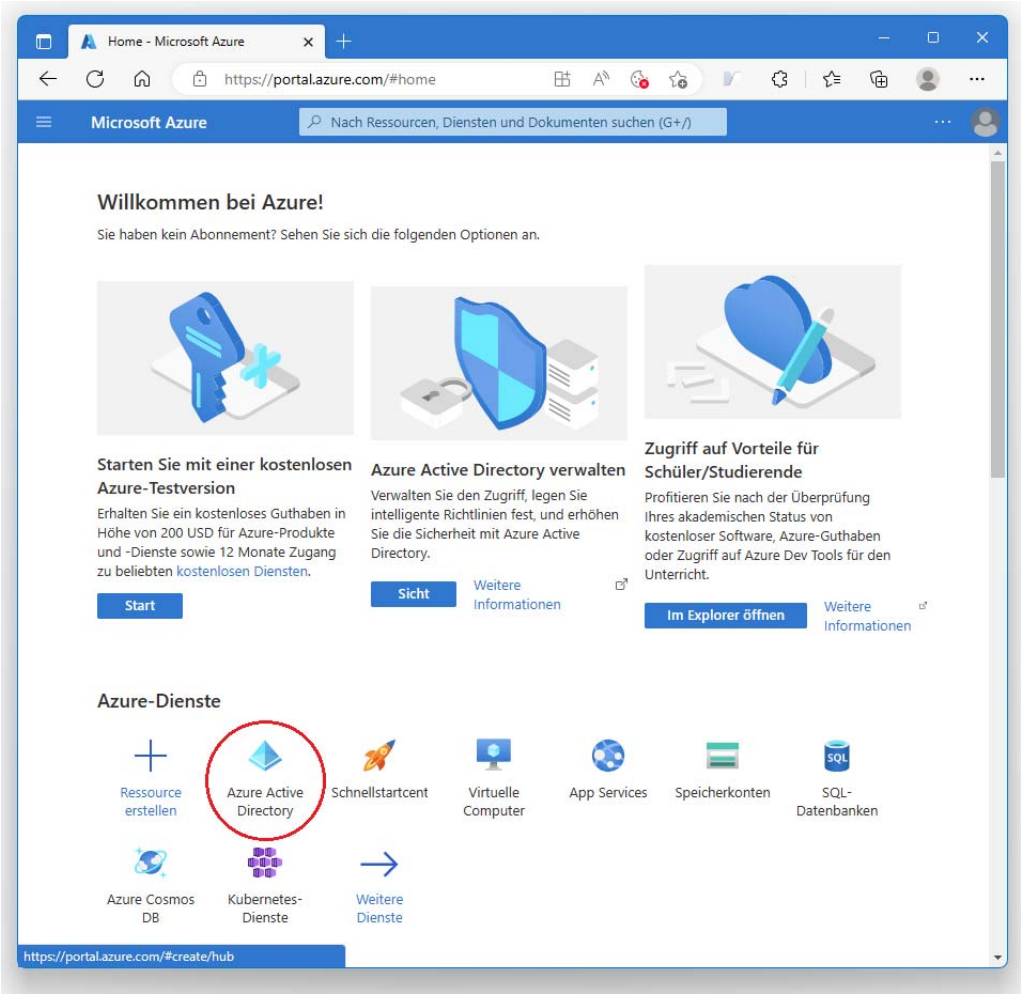
Um die Authentifizierung über den DocumentAssembler für Microsoft 365 verwenden zu können, müssen Sie im Azure Active Directory eine benutzerdefinierte Anwendung erstellen und dieser spezifische Berechtigungen zuordnen.

Achtung: Der Zugriff auf den Outlook365-Postkorb über den DocumentAssembler funktioniert nur mit eingeschalteter 2-Faktor-Authentifizierung.

Wo Sie Ihre Mandanten-ID finden und wie Sie eine Anwendungs-ID für die Benutzung mit dem DocumentAssembler erstellen, wird im weiteren Verlauf erläutert.

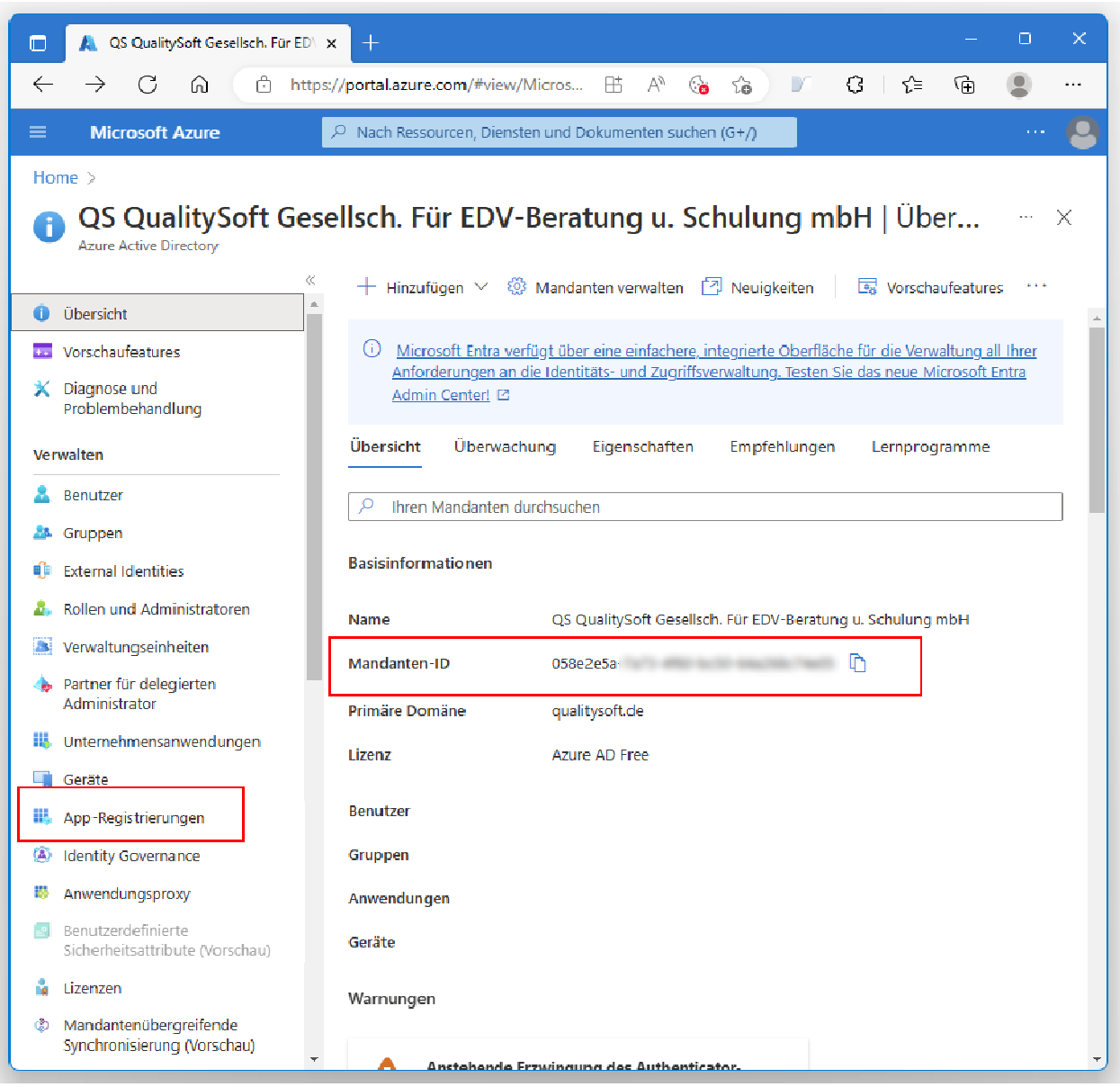
Melden Sie sich am Azure-Portal (https://portal.azure.com/#home) als Administrator an.

Dort wählen Sie unter "Azure-Dienste" das "Azure Active Directory":

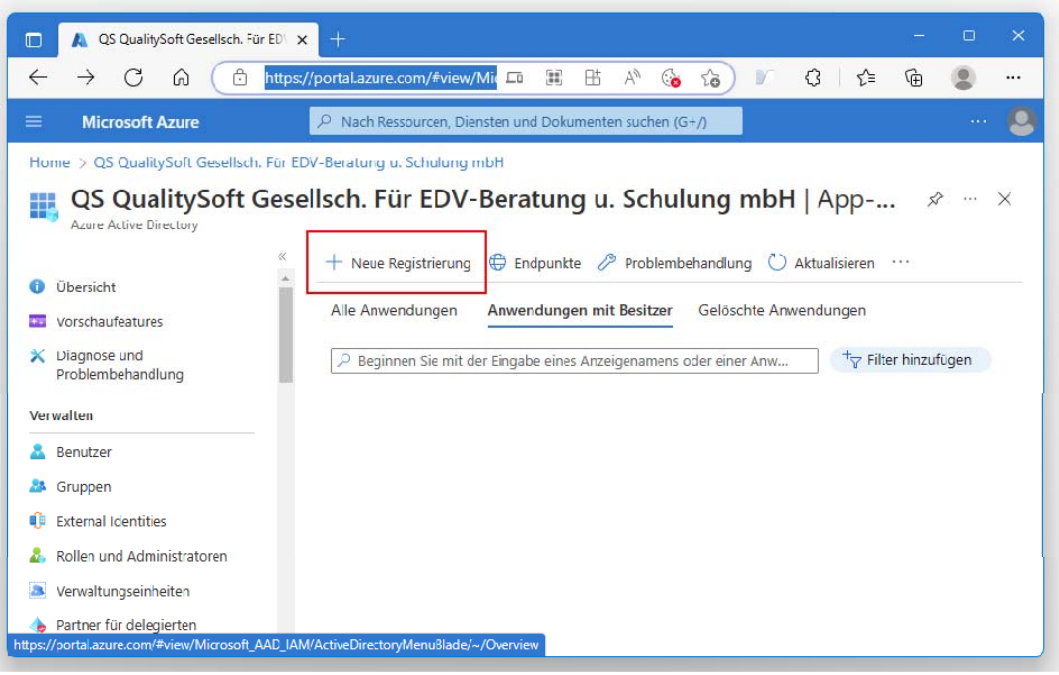


Hier finden Sie auch Ihre Mandanten-ID (Tenant), die Sie für die Konfiguration im DocumenAssembler später benötigen.

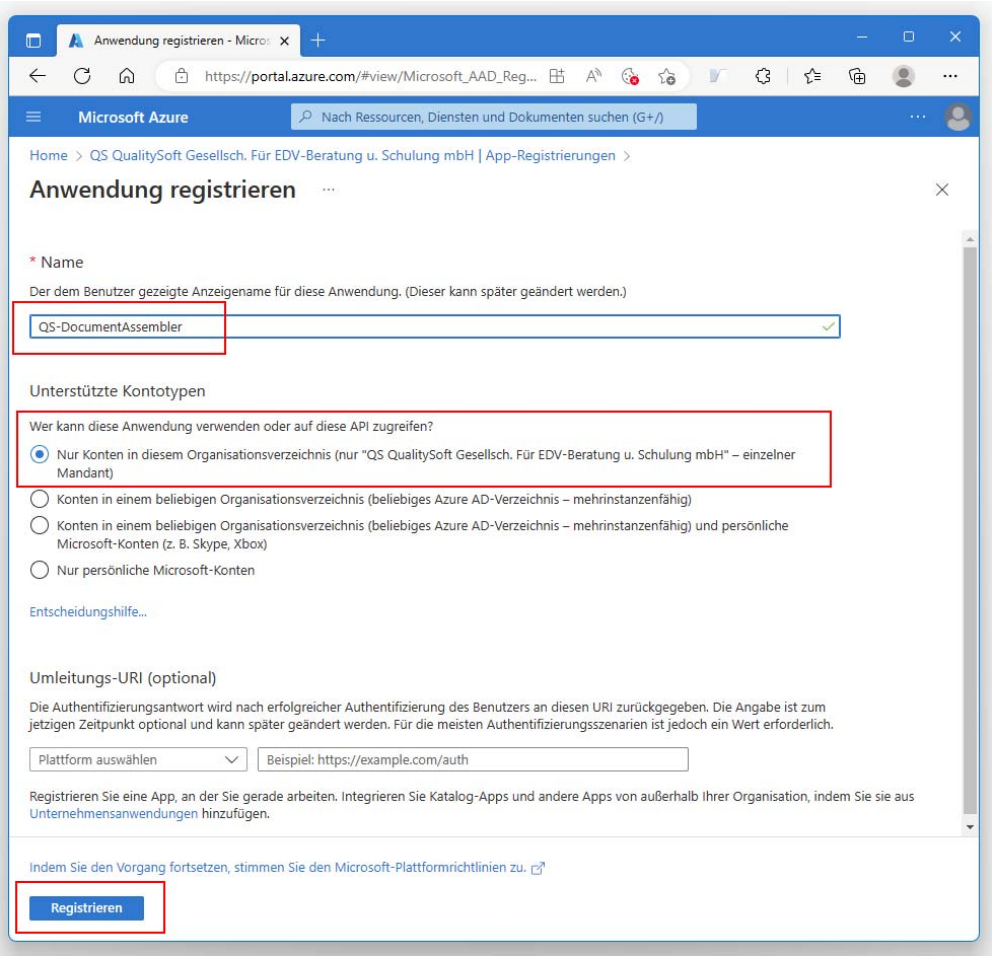
Gehen Sie dann auf "App-Registrierungen"


QS QualitySoft Gesellsch. Für ED
https://portal.azure.com/#view/Micros...
Microsoft Azure
Nach Ressourcen, Diensten und Dokumenten suchen (G+/)
Home
QS QualitySoft Gesellsch. Für EDV-Beratung u. Schulung mbH | Über...
Azure Active Directory
Übersicht
Vorschaufeatures
Diagnose und Problembehandlung
Verwalten
Benutzer
Gruppen
External Identities
Rollen und Administratoren
Verwaltungseinheiten
Partner für delegierten Administrator
Unternehmensanwendungen
Geräte
App-Registrierungen
Identity Governance
Anwendungsproxy
Benutzerdefinierte Sicherheitsattribute (Vorschau)
Lizenzen
Mandantenübergreifende Synchronisierung (Vorschau)
Hinzufügen
Mandanten verwalten
Neuigkeiten
Vorschaufeatures
Microsoft Entra verfügt über eine einfachere, integrierte Oberfläche für die Verwaltung all Ihrer Anforderungen an die Identitäts- und Zugriffsverwaltung. Testen Sie das neue Microsoft Entra Admin Center!
Übersicht
Überwachung
Eigenschaften
Empfehlungen
Lernprogramme
Ihren Mandanten durchsuchen
Basisinformationen
Name
QS QualitySoft Gesellsch. Für EDV-Beratung u. Schulung mbH
Mandanten-ID
058e2e5a
Primäre Domäne
qualitysoft.de
Lizenz
Azure AD Free
Benutzer
Gruppen
Anwendungen
Geräte
Warnungen

Klicken Sie dann auf "Neue Registrierung".

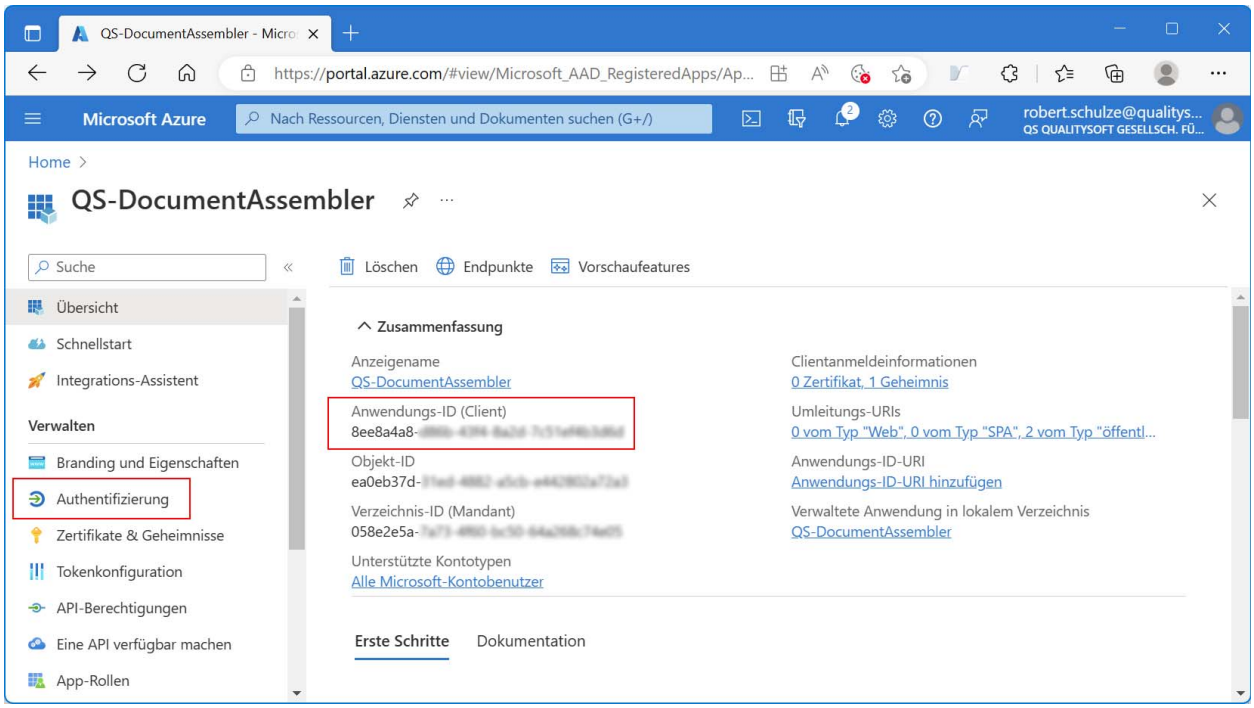


Spezifizieren Sie einen Namen für Ihre benutzerdefinierte Anwendung, z.B. QS-DocumentAssembler und klicken Sie dann auf "Registrieren". Sie können aber auch einen individuellen Namen vergeben.

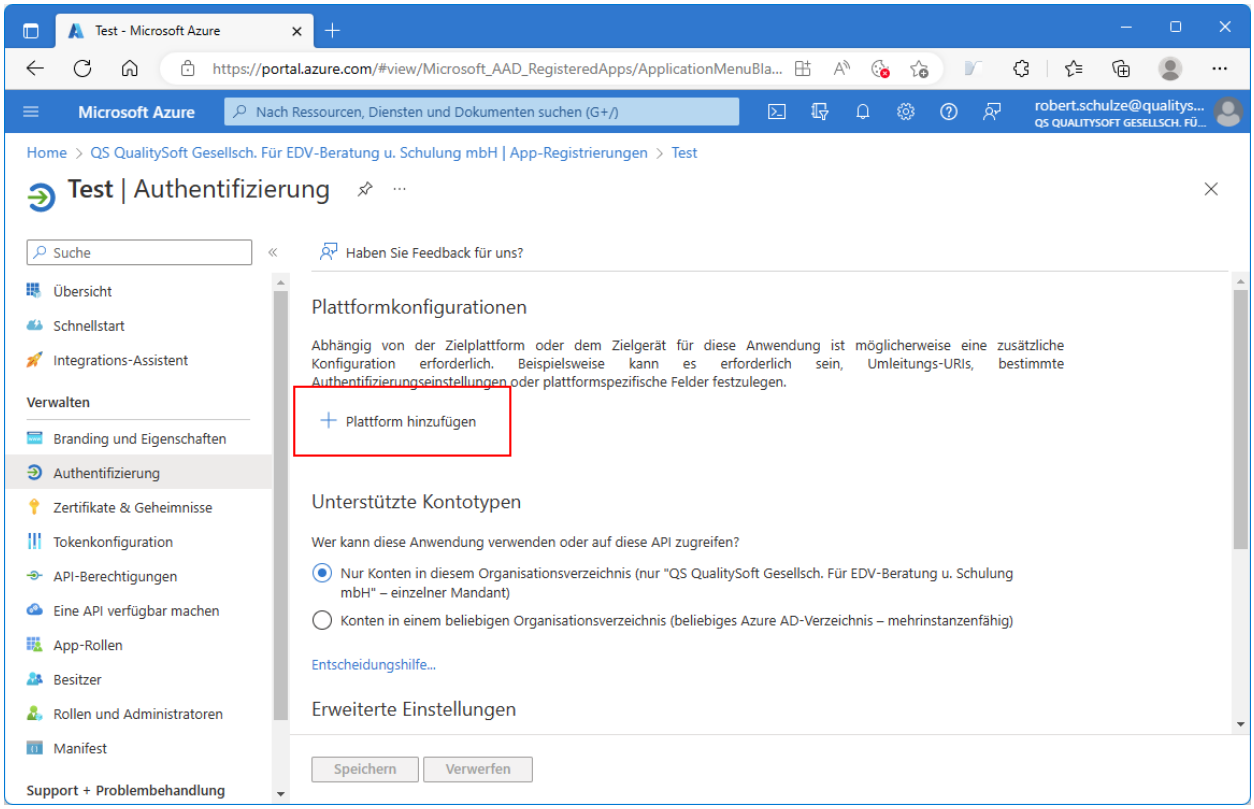


Sie haben jetzt eine App in Azure erstellt. Hier finden Sie jetzt Ihre Anwendungs-ID, die Sie später für die Konfiguration im DocumentAssembler benötigen.

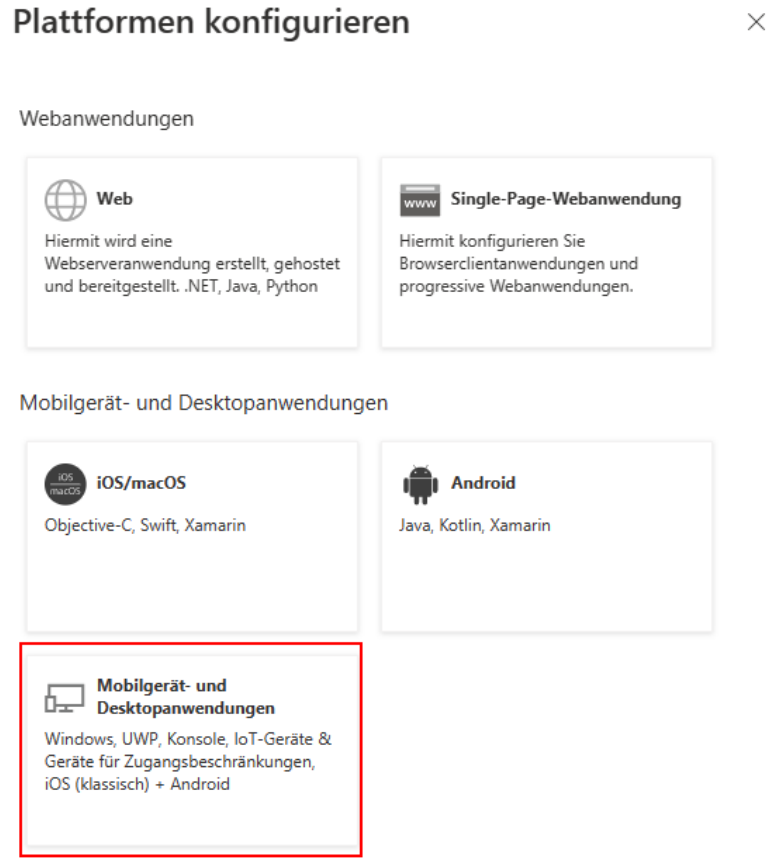
Wählen Sie nun "Authentifizierung".



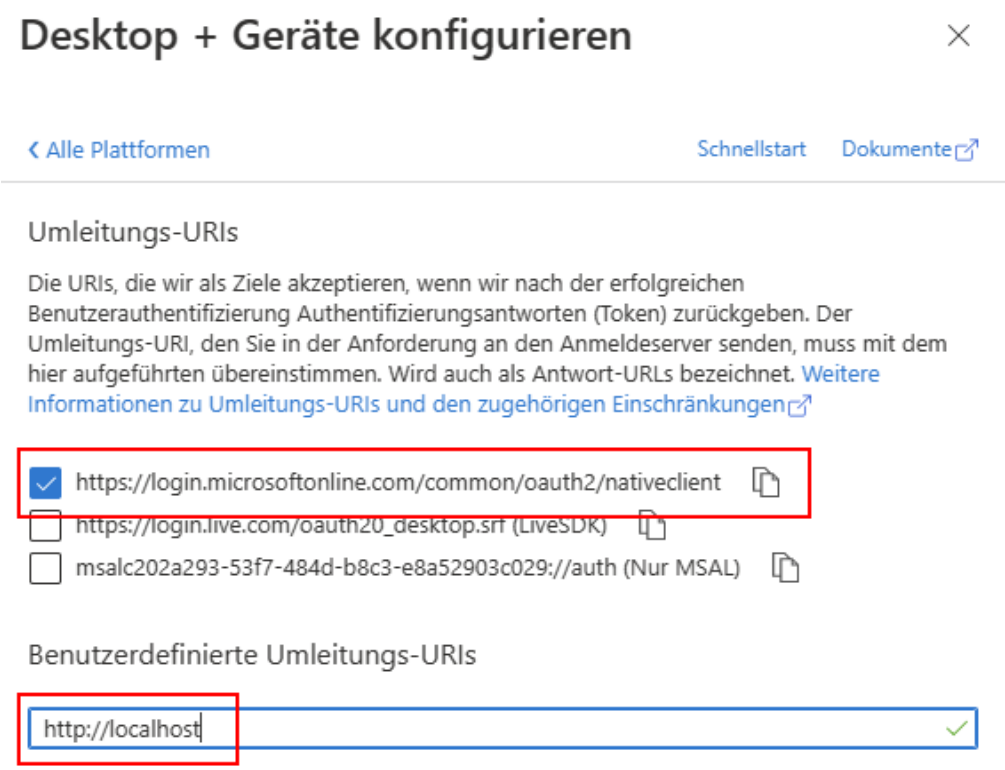
Bei Plattformkonfigurationen wählen Sie jetzt "Plattform hinzufügen.



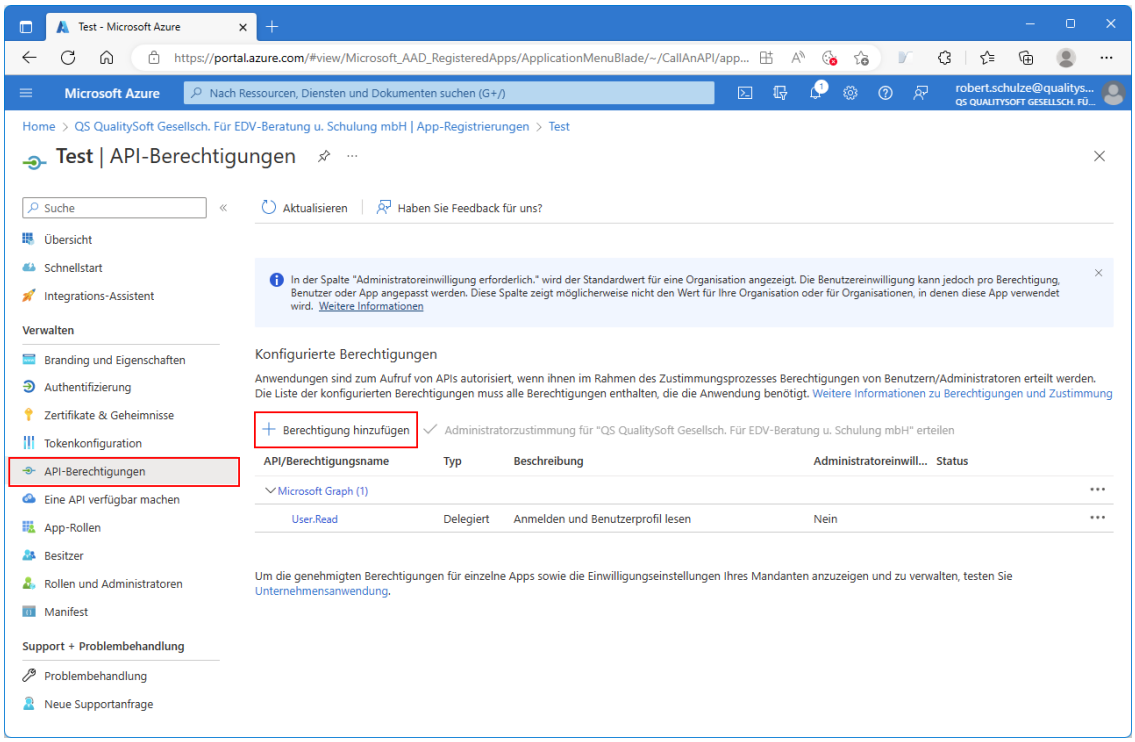
Bei "Plattformen konfigurieren" wählen Sie die Kachel "Mobilgerät und Desktopanwendungen".



Konfigurieren Sie Ihre App wie folgt und speichern dann Ihre Einstellungen:



Jetzt müssen Sie API-Berechtigungen und die Administratorgenehmigung für Ihre Domain erteilen. Dazu klicken Sie auf "API-Berechtigungen" und dann auf "Berechtigung hinzufügen".

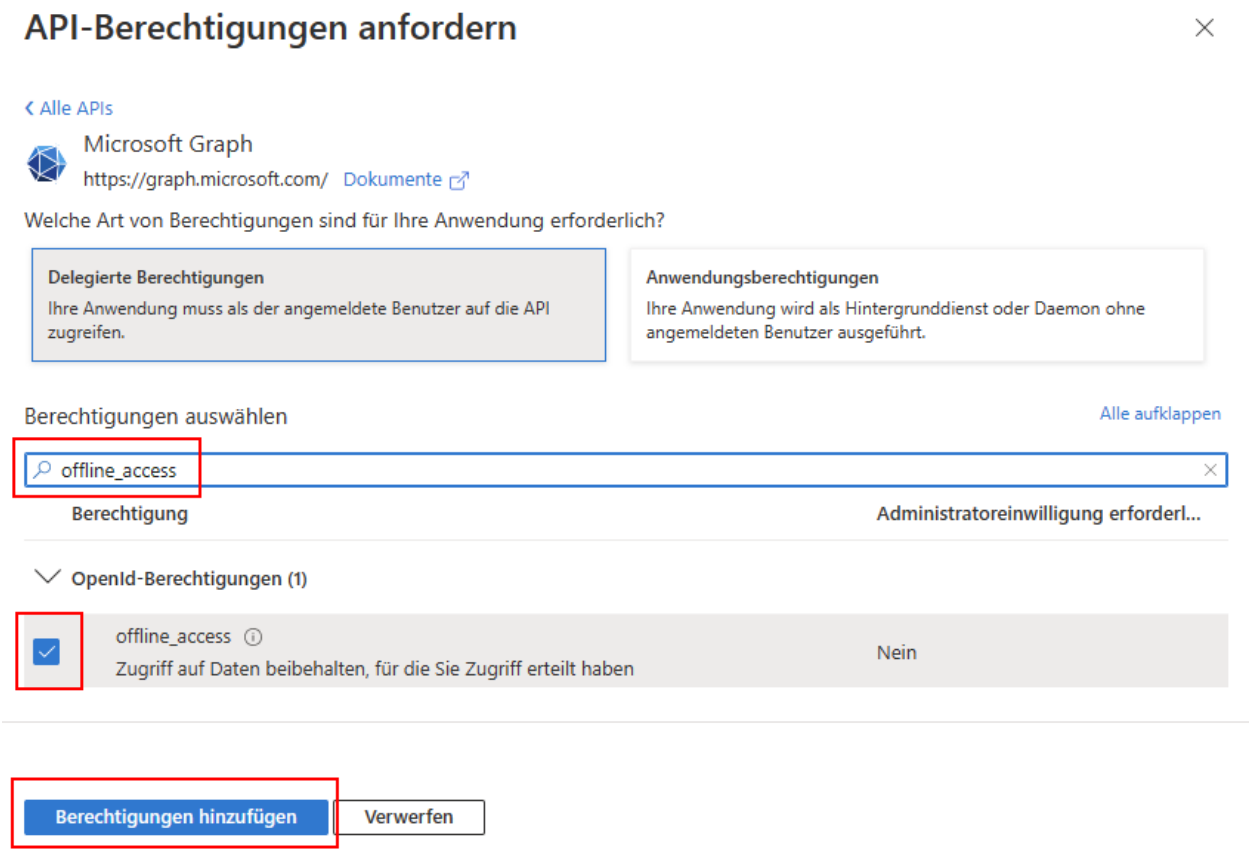


In den API-Berechtigen wählen Sie "Microsoft Graph" und dann "Delegierte Berechtigungen".

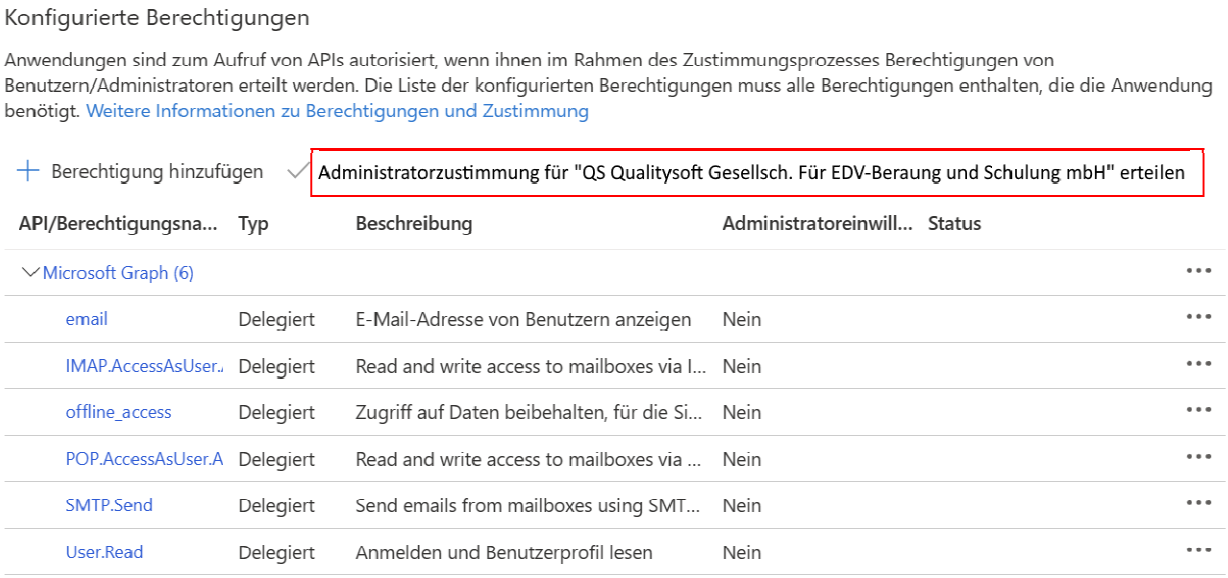
Bei den "Delegierten Berechtigungen" müssen Sie die folgenden Berechtigungsbereiche aktivieren:

offline_access
email
IMAP.AccessAsUser.All
POP.AccessAsUser.All
SMTP.Send

Dazu geben Sie jede Berechtigung in das Suchfeld ein und aktivieren sie anschließend. Wenn Sie alle benötigten Berechtigungen ausgewählt haben, können Sie auf "Berechtigungen hinzufügen" klicken.



Abschließend müssen Sie noch die Administratorzustimmung zu den konfigurierten Berechtigungen erteilen:



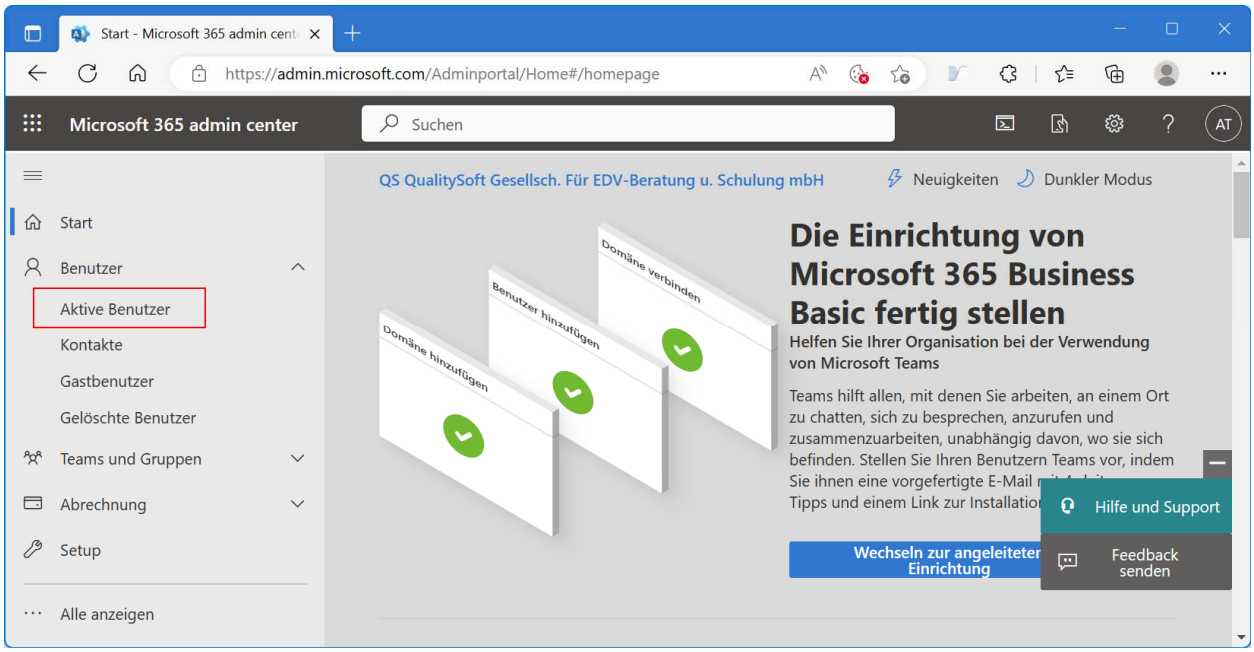
Die Anwendungskonfiguration ist damit abgeschlossen. Sie können jetzt mit der Mandanten-ID und Anwendungs-ID die Konfiguration des Mail-Abrufs im DocumentAssembler fortsetzen.

3.2.2.1.1 SMTP-Freischaltung für das Senden von Mails über Outlook365

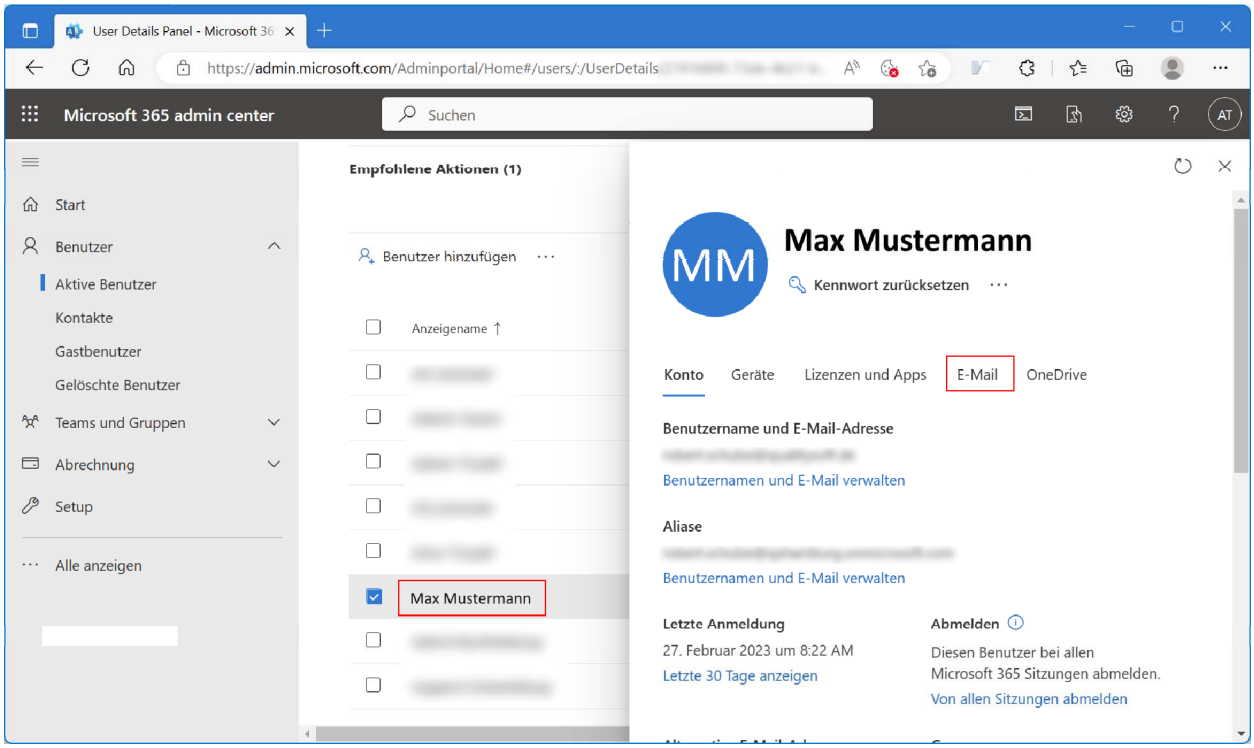
Wenn Sie zusätzlich Mails über Outlook365 aus dem DocumentAssembler heraus versenden wollen, müssen Sie die SMTP-Berechtigung für den entsprechenden Benutzer freischalten.

Verwenden Sie das Microsoft 365 Admin Center zum Aktivieren oder Deaktivieren der SMTP-Authentifizierung für bestimmte Postfächer:

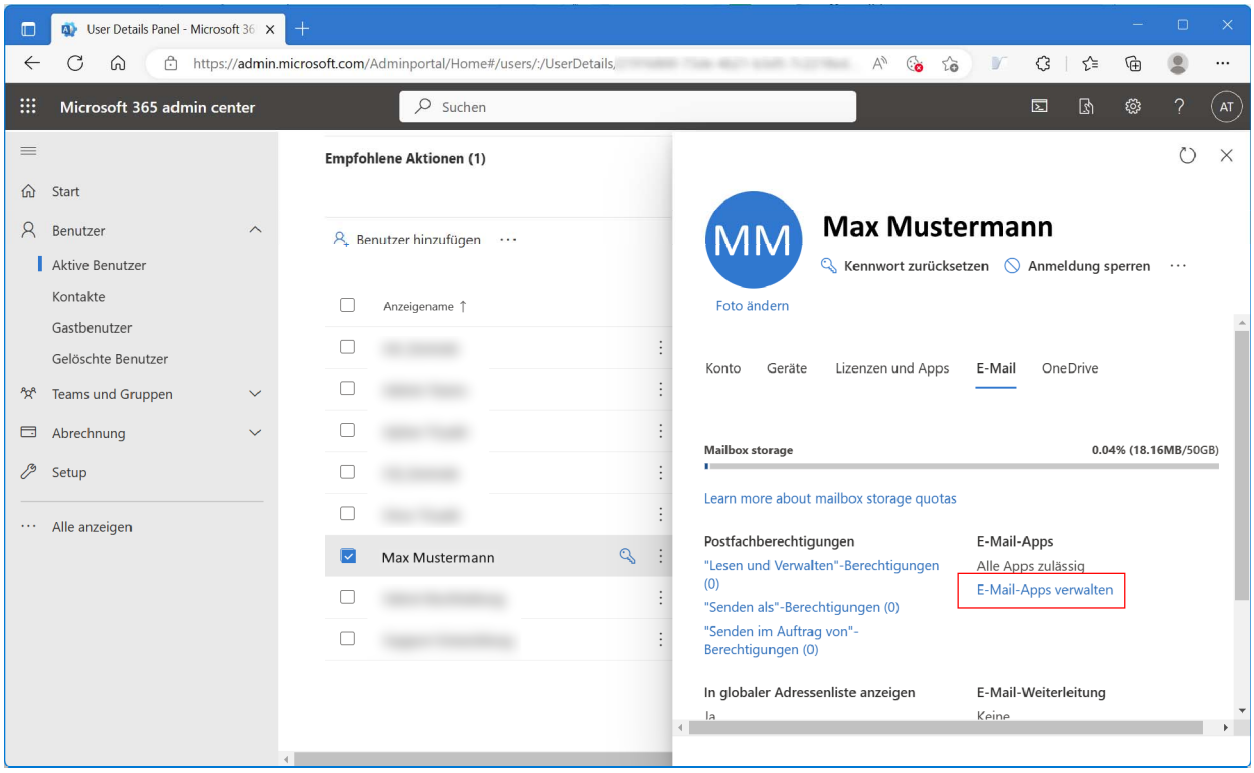
Öffnen Sie das Microsoft 365 Admin Center unter https://admin.microsoft.com/Adminportal/Home und wechseln zu Benutzer > Aktive Benutzer.



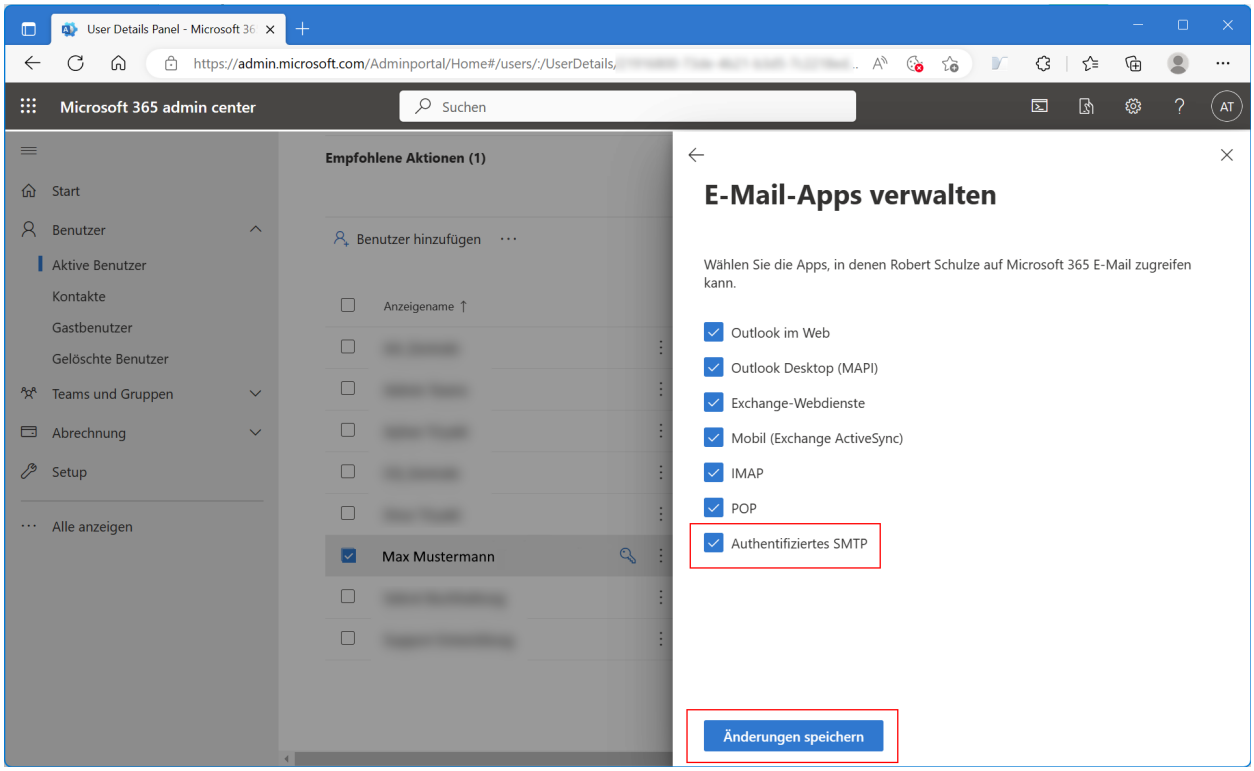
Wählen Sie den Benutzer aus, und klicken Sie im angezeigten Flyout auf E-Mail.



Klicken Sie im Abschnitt E-Mail-Apps auf E-Mail-Apps verwalten.



Überprüfen Sie die Einstellung für "Authentifiziertes SMTP" und wählen Sie es ggf. aus. Speichern Sie anschließend Ihre Änderungen.



Der ausgewählte Benutzer kann jetzt auch Mails über den DocumentAssembler versenden.

3.2.3 Registerkarte "Verarbeitung"

Der QS-DocumentAssembler unterscheidet zwei Verarbeitungsoptionen, den sogenannten Einzelseitenmodus und den Stapelmodus, die sich über den Tab "Barcode/OCR" erreichen lassen.

Einzelseitenmodus: "Jede Seite als Einzelbild behandeln"

Zwischen den einzelnen Quelldateien bzw. Seiten innerhalb einer mehrseitigen Quelldatei besteht kein Zusammenhang. Jede Seite wird für sich behandelt. Wenn kein Barcode auf einer Quellseite gefunden wird, besteht nur die Möglichkeit, diese Seite auszusortieren.

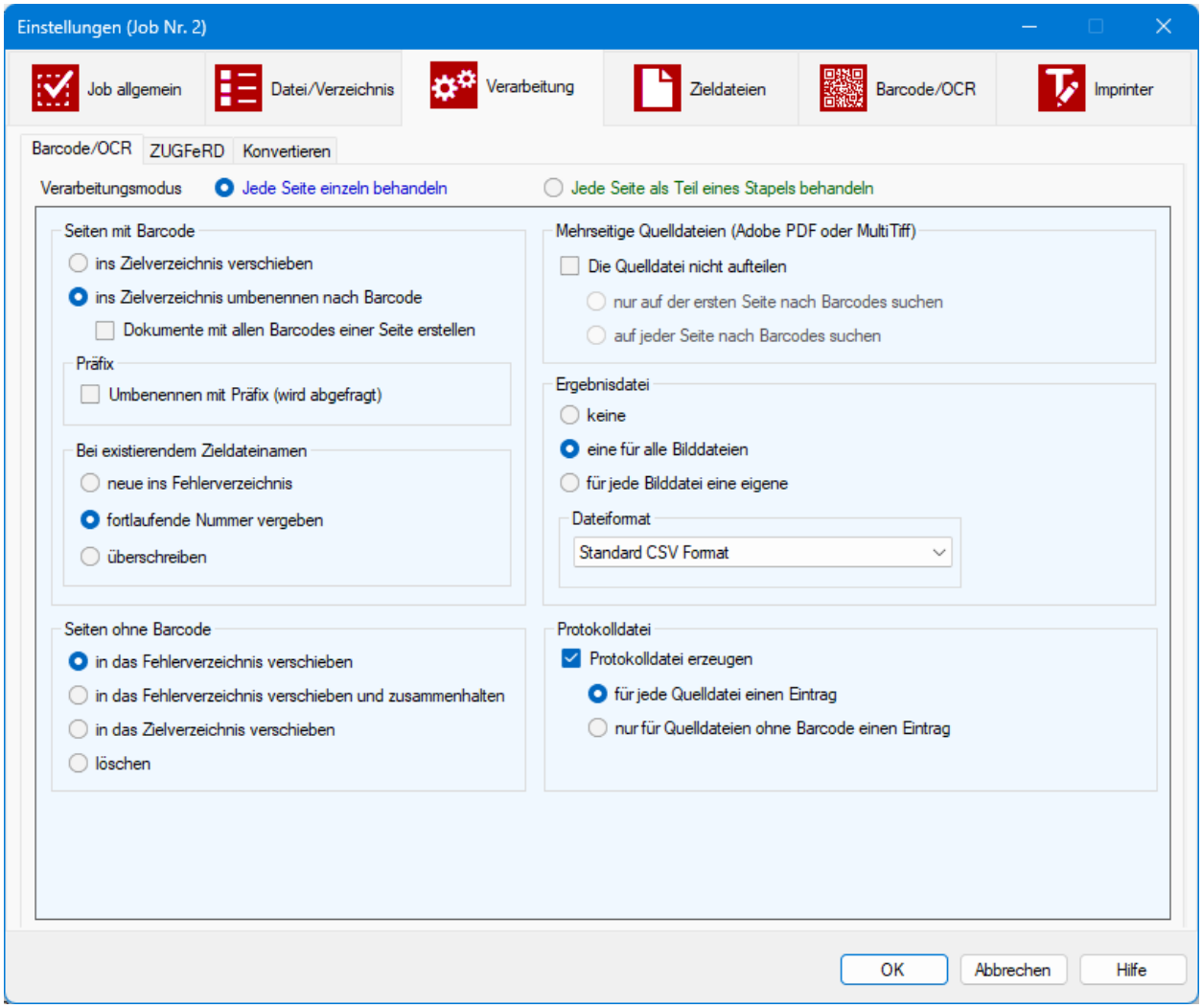
Stapelmodus: "Jede Seite als Teil eines Stapels behandeln"

Ein- oder mehrseitige Quelldokumente werden als Stapel bzw. Teil eines Stapels betrachtet. Mehrere Quelldokumente können zusammengehören, zumindest die erste Seite eines Stapels ist mit einem Barcode versehen. Der Wechsel des Barcodeinhalts impliziert den Beginn eines neuen Stapels. Seiten ohne Barcode können als Folgeseiten dem Stapel zugeordnet werden.

Diese zwei Verarbeitungsmodi werden durch einen BLAUEN und einen GRÜNEN Bereich innerhalb der Registerkarte repräsentiert.

Außerdem unterstützt der DocumentAssembler die Verarbeitung von ZUGFeRD-Dokumenten sowie die einfache Umwandlung von Quelldokumenten in ein anderes Dokumentenformat ohne Barcodelesung. Dies lässt sich jeweils über den über den Tab "ZUGFeRD" bzw. den Tab "Konvertieren" konfigurieren.

3.2.3.1 Verarbeitung Einzelseitenmodus (BLAU)



Seiten mit Barcode

Diese Seiten werden immer ins Zielverzeichnis verschoben. Wenn gewünscht, kann das Dokument als Namen den erkannten Barcodes bekommen.

"ins Zielverzeichnis verschieben"
Die Zieldatei erhält den Namen der Quelldatei plus Seitennummer.

"ins Zielverzeichnis umbenennen nach Barcode"
Die Zieldatei erhält den Namen des erkannten Barcodes.

"Dokumente mit allen Barcodes einer Seite erstellen"
Befinden sich mehrere Barcodes auf einer Seite, so können auch mehrere Zieldokumente mit den gefundenen Barcodes erstellt werden.

Präfix

Die Dateinamen der Zieldateien können zusätzlich noch mit einem Präfix versehen werden. Dieses Präfix wird dann zu Beginn der Verarbeitung in einem Dialog abgefragt. Es kann z. B. genutzt werden, um Abteilungskennungen oder Dokumentkategorie-Kennungen (RE=Rechnung, LI=Lieferschein) zu vergeben, die nicht im Barcode enthalten sind.

Seiten ohne Barcode

Hier wird gesteuert was passieren soll, wenn kein Barcode gefunden wurde.

"in das Fehlerverzeichnis verschieben"
: Die Zieldatei wird in das Fehlerverzeichnis geschrieben und erhält den Namen der Quelldatei plus Seitennummer

"in das Fehlerverzeichnis verschieben und zusammenhalten"
: Die fehlerhafte Seite wird unter dem Namen der Quelldatei in das Fehlerverzeichnis geschrieben. Bei mehrseitigen Quelldateien werden fehlerhafte Seiten aus der Quelldatei an die Datei im Fehlerverzeichnis angehängt, so dass eine einzige Fehlerdatei mit allen fehlerhaften Seiten erstellt wird.

"in das Zielverzeichnisverschieben"
: Die Zieldatei wird in das Zielverzeichnis geschrieben und erhält den Namen der Quelldatei plus Seitennummer

"Löschen"
: Es wird keine Zieldatei geschrieben.

Bitte beachten Sie beim Löschen, dass ein Barcode nicht erkannt werden muss, obwohl ein Barcode auf der Seite aufgebracht ist. Dies kann passieren, wenn das Bildmaterial stark verschmutzt ist oder die Barcode-Einstellungen nicht zu dem Barcode auf dem Bild passen.

Bei existierendem Dateinamen

Hier wird gesteuert was passieren soll, wenn im Zielverzeichnis bereits eine Datei gleichen Namens vorhanden ist.

"neue ins Fehlerverzeichnis"
: Wenn der Zieldateiname existiert, wird die Seite ins Fehlerverzeichnis mit lfd. Nummer geschrieben.

"fortlaufende Nummer vergeben
: Bei der Option "fortlaufende Nummer vergeben" bilden sich die Dateinamen immer aus <BARCODE>.NNNN.<DATEIENDUNG>. Bereits die erste Zieldatei erhält schon eine fortlaufende Nummer NNNN=0001 im Dateinamen.

"überschreiben"
: Die im Zielverzeichnis existierende Datei wird überschrieben.

Mehrseitige Quelldateien (Adobe PDF oder Multipage TIFF)

Wenn die zu verarbeitenden Quelldateien mehrseitige Dateien sind, kann hier das Verarbeitungsverhalten der Dateien im Einzelmodus gesteuert werden:

Beim Verschieben, Umbenennen usw. wird die Datei als Einheit behandelt und nicht in einzelne Seiten getrennt. Im Sonderfall kann, wenn auf der ersten Seite kein Barcode gefunden wurde, auf den Folgeseiten in der Datei nach Barcodes gesucht werden.

"Die Quelldatei nicht aufteilen"
: Das Quelldokument wird nicht in einzelne Seiten aufgeteilt, sondern als eine Einheit betrachtet, und zwar unabhängig davon ob ein Barcode gefunden wird oder nicht.

"*nur auf der ersten Seite nach Barcodes suchen*"
: Es wird lediglich auf der ersten Seite nach Barcodes gesucht. Wird kein Barcode gefunden, wird die definierte Fehlerbehandlung ausgeführt.

"*auf jeder Seite nach Barcodes suchen*"

Wenn auf der ersten Seite kein Barcode gefunden wurde, wird auf den Folgeseiten solange nach einem Barcode gesucht, bis der erste Barcode gefunden wird. Danach wird die Suche sofort abgebrochen. Wurde ein Barcode gefunden ist die weitere Verarbeitung so, als ob dieser Barcode auf der ersten Seite gefunden wurde. Wenn kein Barcode gefunden wurde, wird die definierte Fehlerbehandlung ausgeführt.

Ergebnisdatei

Es kann eingestellt werden, ob eine Ergebnisdatei mit den erkannten Barcodes erzeugt werden soll oder nicht. Wenn eine gemeinsame Ergebnisdatei für alle Bilddateien erzeugt werden soll, ergibt sich der Dateiname aus dem in der Registerkarte "Datei/Verzeichnis " angegebenen Dateinamen. Wenn pro Bilddatei eine einzelne Ergebnisdatei erzeugt werden soll, erhält die Datei immer denselben Namen wie die zugehörige Zieldatei, mit der Dateiendung ".txt".
Es kann zwischen unterschiedlichen Exportformaten gewählt werden.

"eine für alle Bilddateien"

- *Standard CSV Format:*
  "BARCODE";"FILENAME"
  "123456789";"c:\docasm\ziel\123456789.0001.tif"

- *Alle Ergebnisse, Semikolon getrennt:*
  c:\docasm\ziel\dok01.tif;123456789;ABCD;XYZ

- *Alle Ergebnisse, Semikolon getrennt, kurz:*
  dok01.tif;123456789;ABCD;XYZ

"für jede Bilddatei eine eigene"

- *Standard mit Ergebnis-Header pro Ergebnis:*
  >>> RESULT: 1
  >>>
  123456789
  >>>

- *Standard ANSI, max. ein Ergebnis:*
  123456789

- *Standard UTF8, max. ein Ergebnis:*
  123456789

- *Alle Ergebnisse, Semikolon getrennt:*
  c:\docasm\ziel\ dok01.tif;123456789;ABCD;XYZ

- *Alle Ergebnisse, Semikolon getrennt, kurz:*
  dok01.tif;123456789;ABCD;XYZ

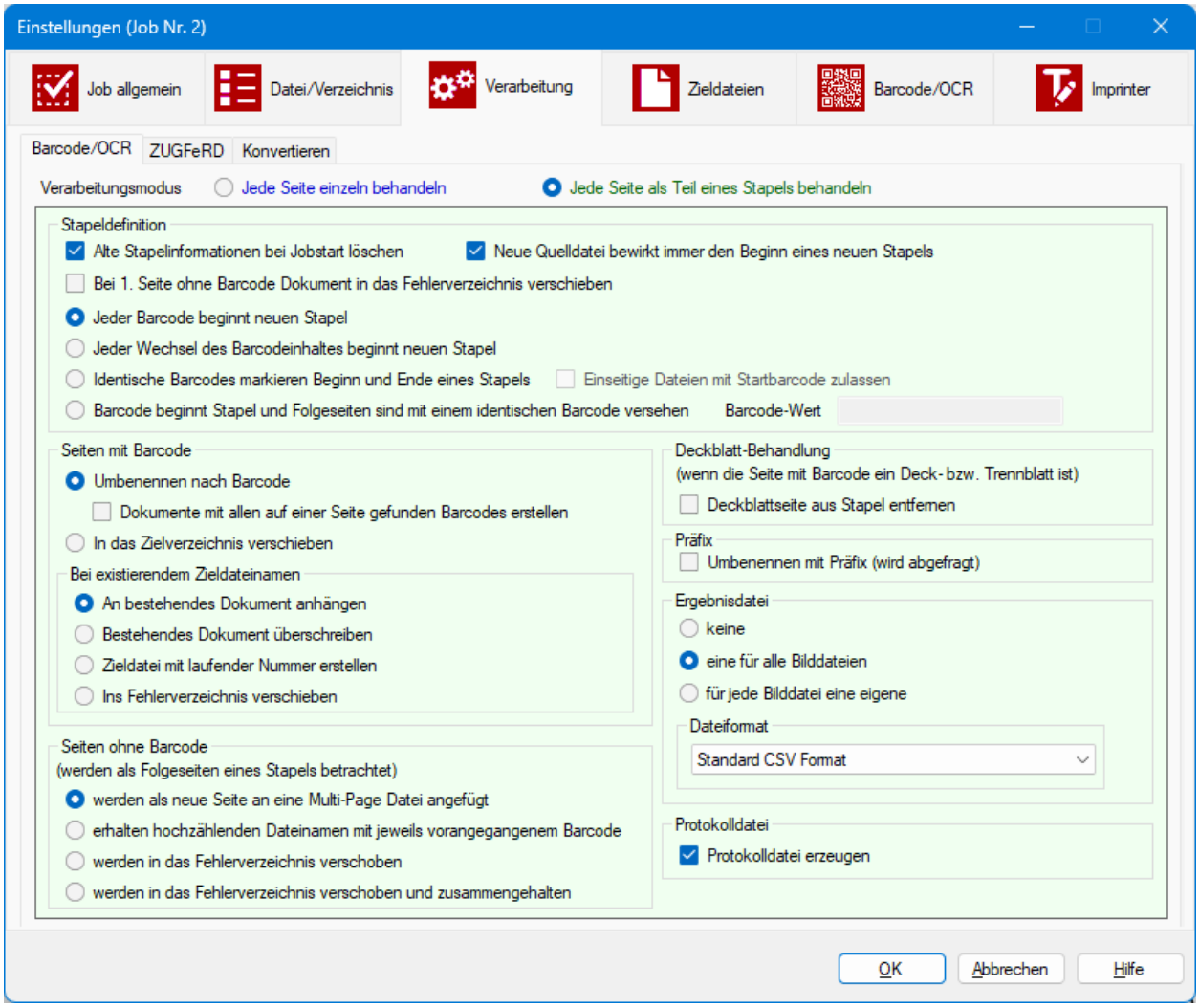
- *Alle Ergebnisse mit Positionsinfo:*
  BARCODE;X;Y;DX;DY;ORIENTATION

123456789;155;288;389;123;1

Protokolldatei

In der Protokolldatei wird die Verarbeitung jeder einzelnen Seite dokumentiert. Es kann zudem konfiguriert werden, ob alle oder nur Seiten ohne Barcode protokolliert werden sollen. Dies kann dazu genutzt werden, eine Fehlerliste zu erzeugen. (siehe Anhang Kapitel 7.1, Protokolldatei)

3.2.3.2 Verarbeitung Stapelmodus (GRÜN)



Stapeldefinition

"*Alte Stapelinformationen bei Jobstart löschen*"

Die erste Seite des Dokuments nach Start eines Jobs kann, wenn es ohne Barcode ist, als Folgeseite eines vorangegangenen Stapels behandelt werden. Dadurch wird die Seite an das letzte erzeugte Dokument des vorherigen Durchlaufs angehängt. Wenn alle Stapelinformationen beim Programmstart gelöscht werden wird verhindert, dass neue Dokumente einem "alten" Stapel zugeordnet werden.

"Neue Quelldatei bewirkt im immer den Beginn eines neuen Stapels"

Diese Option stellt sicher, dass bei einer neuen Datei auf jeden Fall ein neuer Stapel begonnen wird. Ohne Barcode auf der ersten Seite wird diese Seite dann als fehlerhaft verarbeitet.

"Bei 1. Seite ohne Barcode Dokument in das Fehlerverzeichnis verschieben"

Wenn auf der 1. Seite eines Quelldokumentes kein Barcode gefunden wird, wird die ganze Datei in das Fehlerverzeichnis verschoben.

Jeder gefundene Barcode definiert den Beginn eines neuen Stapels

Bei dieser Konstellation wird, abweichend von der vorherigen Definition, bei jedem gefundenen Barcode ein neuer Stapel definiert. Dies dient u.a. dazu, im Stapelmodus Dokumente mit identischen Barcodes zu trennen anstatt sie in eine Zieldatei zu schreiben.

Jeder Barcodewechsel definiert den Beginn eines neuen Stapels

Wenn eine Seite mit Barcode gefunden wird, wird ein neuer Stapel erstellt. Alle folgenden Seiten mit und ohne Barcode werden an diesen Stapel angehängt bis ein neuer vom vorherigen Barcode abweichender Barcode gefunden wird, der wiederum einen neuen Stapel beginnt.

Identische Barcodes markieren Beginn und Ende eines Stapels

Es besteht immer die Gefahr, dass Barcodes nicht erkannt werden und dann zwei Stapel zu einem zusammengefasst werden. Um dies zu verhindern, kann man den Beginn und das Ende eines Stapels mit einem Identischen Barcode zu kennzeichnen. Wenn der zweite Barcode hier nicht gefunden werden kann, wird der gesamte Bereich bis zum nächsten gefundenen Barcode als Dokument im Fehlerverzeichnis gespeichert.

"Einseitige Dateien mit Startbarcode zulassen "

Wenn ein Stapel nur aus einer Seite besteht, dann soll auf den identischen Barcode als Zeichen des Endes des Stapels verzichtet werden. Einseitige Dateien brauchen also nur einen Startbarcode zu haben.

Barcode beginnt Stapel und Folgeseiten sind mit identischem Barcode versehen

Es besteht immer die Gefahr, dass Barcodes nicht erkannt werden und dann zwei Stapel zu einem zusammengefasst werden. Um das zu verhindern können Sie Folgeseiten manuell mit einem identischen Barcode ausstatten.

"Barcode-Wert "

Wenn man auf jeder Folgeseite einen Platzhalter-Barcode mit einem festen, gleichbleibenden Wert anbringt, kann der QS-DocumentAssembler unterscheiden:
- Gar kein Barcode auf der Seite: Fehler
- Definierter Barcode auf der Seite: Folgeseite
- Anderer Barcode auf der Seite: Neuer Stapel beginnt

Das fehlerhafte Zusammenfassen von zwei Stapeln kann so verhindert werden.
Geben Sie bitte den spezifischen Wert für den Barcode an.

Seiten mit Barcode

"Umbenennen nach Barcode"

Bei gefundenem Barcode wird dieser Barcode für die Benennung der Zieldatei übernommen.

Kommt ein Barcode im Quelldokument im Wechsel mit anderen Barcodes mehrfach vor, werden diese Seiten an eventuell vorhandene Zieldokumente angehängt. Um dies zu verhindern, kann die Zieldatei mit einer lfd. Nummer versehen werden. Bei wiederkehrenden identischen Barcodes in einem Quelldokument würde dann die lfd. Nummer erhöht werden.
Zusätzlich kann man hier Zieldateien mit allen auf einer Seite gefundenen Barcodes erstellen.

Dokumente mit allen auf der Seite gefundenen Barcodes erstellen

Wenn auf einer Seite mehrere Barcodes gefunden werden, wird die Quelle für jeden Barcode in dem Zielverzeichnis vervielfacht.

"In das Zielverzeichnis verschieben"
Bei gefundenem Barcode wird der Quelldateiname für die Benennung der Zieldatei übernommen.

Seiten ohne Barcode

Wenn auf einer Seite kein Barcode gefunden wurde, wird diese Seite immer als Folgeseite eines bestehenden Stapels betrachtet. Ist die Seite die erste Seite einer Quelldatei und kann diese Seite keinem vorherigen Stapel (durch Löschen der Stapelinformationen) zugeordnet werden, wird diese Seite ins Fehlerverzeichnis verschoben.

"werden als neue Seite an eine Multipage-Datei angefügt"
Seiten ohne Barcode werden als Folgeseite an eine bestehende Datei angehängt. Der Dateiname ergibt sich dabei aus der Behandlung von Seiten mit erkanntem Barcode.

"erhalten hochzählenden Dateinamen mit jeweils vorangegangenem Barcode"
Die Folgeseiten ohne Barcode werden nach dem zuletzt gelesenen Barcode umbenannt und bekommen zusätzlich eine lfd. Nummer.

"werden in das Fehlerverzeichnis verschoben."
Seiten ohne Barcode werden in das Fehlerverzeichnis verschoben.

"werden in das Fehlerverzeichnis verschoben und zusammengehalten"
Die fehlerhafte Seite wird unter dem Namen der Quelldatei in das Fehlerverzeichnis geschrieben. Bei mehrseitigen Quelldateien werden fehlerhafte Seiten aus der Quelldatei an die Datei im Fehlerverzeichnis angehängt, so dass eine einzige Fehlerdatei mit allen fehlerhaften Seiten erstellt wird.

Bei existierendem Zieldateinamen

Entgegen dem bisherigen Verhalten, bei existierendem Dateinamen Seiten einfach an das bestehende Dokument anzuhängen, können Sie jetzt das Konfliktverhalten bei bestehendem Dateinamen selber bestimmen:

"An bestehendes Dokument anhängen."
Bei Barcodewiederholung werden Seiten an das bestehende Dokument angehängt.

"Bestehendes Dokument überschreiben."
Bei Barcodewiederholung wird das bestehende Dokument überschrieben.

"Zieldatei mit laufender Nummer erstellen."
Bei Barcodewiederholung wird ein Zieldokument mit laufender Nummer erzeugt. Alle Folgeseiten werden an dieses Dokument angehängt, bis der Stapel wechselt.

"Ins Fehlerverzeichnis verschieben."

Bei Barcodewiederholung wird das neue Dokument in das Fehlerverzeichnis verschoben.

Ergebnisdatei

Es kann eingestellt werden, ob eine Ergebnisdatei mit den erkannten Barcodes erzeugt werden soll. Wenn eine gemeinsame Ergebnisdatei für alle Bilddateien erzeugt werden soll, ergibt sich der Dateiname aus der in der Registerkarte "Datei/Verzeichnis " konfigurierten Ergebnisdatei. Wenn pro Bilddatei eine einzelne Ergebnisdatei erzeugt werden soll, erhält die Datei immer denselben Namen wie das zugehörige Bild mit der Dateiendung ".txt".
Es kann zwischen unterschiedlichen Exportformaten gewählt werden.

"eine für alle Bilddateien"

- *Standard CSV Format:*
  "BARCODE";"FILENAME"
  "123456789";"c:\docasm\ziel\123456789.0001.tif"

- *Alle Ergebnisse, Semikolon getrennt:*
  c:\docasm\ziel\ dok01.tif;123456789;ABCD;XYZ

- *Alle Ergebnisse, Semikolon getrennt, kurz:*
  dok01.tif;123456789;ABCD;XYZ

"für jede Bilddatei eine eigene"

- *Standard mit Ergebnis-Header pro Ergebnis:*
  >>> RESULT: 1
  >>>
  123456789
  >>>

- *Standard ANSI, max. ein Ergebnis:*
  123456789

- *Standard UTF8, max. ein Ergebnis:*
  123456789

- *Alle Ergebnisse, Semikolon getrennt:*
  c:\docasm\ziel\ dok01.tif;123456789;ABCD;XYZ

- *Alle Ergebnisse, Semikolon getrennt, kurz:*
  dok01.tif;123456789;ABCD;XYZ

- *Alle Ergebnisse mit Positionsinfo:*
  BARCODE;X;Y;DX;DY;ORIENTATION

123456789;155;288;389;123;1

"*Protokolldatei*"

In der Protokolldatei wird die Verarbeitung jeder einzelnen Seite dokumentiert. Es kann zudem konfiguriert werden, ob alle oder nur Bilddateien ohne Barcode protokolliert werden sollen. Dies kann dazu genutzt werden, eine Fehlerliste zu erzeugen.

Deckblatt-Behandlung

Wenn man Dokumente ohne Barcode hat und die Stapeltrennung mit nachträglich generierten Barcodes erfolgt, können diese im Rahmen der Verarbeitung entfernt werden.

"Deckblattseite aus Stapel entfernen"

Wenn auf einer Seite ein Barcode erkannt wird, wird dieser Barcode als Stapelname betrachtet und die Seite wird aus der Quelldatei gelöscht.

3.2.3.3 Die Benennung der Zieldateien in der Verarbeitung von Einzelseiten- und Stapelmodus

Die Zieldateien werden je nach Verarbeitungsmodus und ausgewählter Verarbeitungsart unterschiedlich benannt. Im Einzelseitenmodus kann es bis zu einer zweistufigen Nummerierung der Zieldatei kommen, im Stapelmodus bis maximal zu einer einstufigen Benennung. Das Format der Zieldateien entnehmen Sie bitten den folgenden Tabellen.

Einzelseitenmodus

Seiten mit Barcode	**Format Zielverzeichnis**	**Format Fehlerverzeichnis**
Verschieben, neue ins Fehlerverzeichnis	Quelldateiname.tif	Quelldateiname.[Seite].[lfd.Nr].tif
Verschieben, fortlaufende Nummer	Quelldateiname.[Seite].[lfd.Nr].tif	-
Verschieben, überschreiben	Quelldateiname.tif	-
Umbenennen, neue ins Fehlerverzeichnis	Barcode.tif	Barcode.[lfd.Nr].tif
Umbenennen, fortlaufende Nummer	Barcode.[lfd.Nr].tif	-
Umbenennen, überschreiben	Barcode.tif	-
Seiten ohne Barcode		
Zielverzeichnis	Quelldateiname.[Seite].[lfd.Nr].tif	
Fehlerverzeichnis	-	Quelldateiname.[Seite].[lfd.Nr].tif
Löschen	-	-

Die Nummerierung [lfd.Nr] gibt die Anzahl der Vorkommen im Zielverzeichnis an, die Seitenangabe [Seite] bezieht sich auf die Seitennummer im Quelldokument.

Stapelmodus

Seiten mit Barcode	**Format Zielverzeichnis**	**Format Fehlerverzeichnis**
Verschieben ins Zielverzeichnis	Quelldateiname.tif	-
Umbenennen ins Zielverzeichnis	Barcode.tif	-
Umbenennen ins Zielverzeichnis mit fortlaufenden Nummer	Barcode.[lfd.Nr].tif	-
Verschieben ins Zielverzeichnis mit fortlaufender Nummer	Quelldateiname.[lfd.Nr].tif	
Seiten ohne Barcode		-
An Multipage-Datei anhängen	Letzter Zieldateiname	
Vorheriger Barcode mit laufender Nummer	Barcode.[lfd.Nr].tif	-
Erste Seite einer Quelldatei und kein Stapel		Quelldateiname.[lfd.Nr].tif
Bei 1. Seite ohne Barcode Dokument ins Fehlerverzeichnis		Quelldateiname.tif

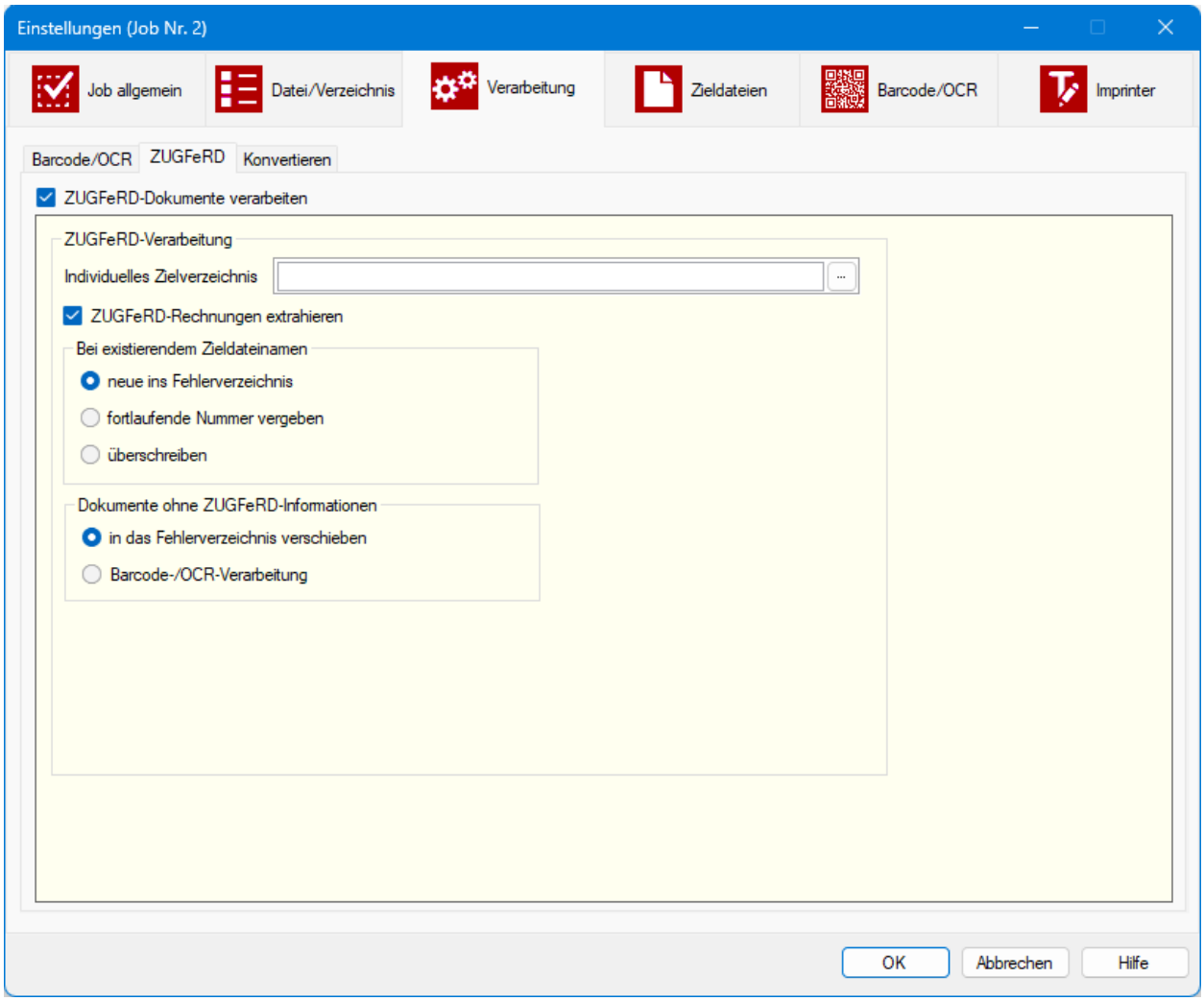
Im Stapelmodus wird bei Vorhandensein eines Dokumentes gleichen Namens grundsätzlich an dieses Dokument angehängt, es sei denn, für das Zielverzeichnis ist explizit eine Nummerierung ausgewählt.

Ist eine Nummerierung ausgewählt, wird die laufende Nummer fortlaufend hochgezählt.

3.2.3.4 ZUGFeRD-Verarbeitung

Der DocumentAssembler kann neben Barcode- und OCR-Verarbeitung auch ZUGFeRD-Dokumente erkennen und verarbeiten. Es besteht die Möglichkeit, entweder nur ZUGFeRD-Dokumente zu erkennen und zu verschieben oder zusätzlich die XML-Rechnungen aus den Dokumenten zu extrahieren.

Grundsätzlich besteht die Möglichkeit, die ZUGFeRD-Verarbeitung separat auszuführen oder mit Barcode-/OCR-Verarbeitung gemischt auszuführen. Bei der gemischten Ausführung werden dann alle nicht ZUGFeRD-Dokumente in die Barcodeverarbeitung überführt.



"*ZUGFeRD-Dokumente verarbeiten*"

Die Verarbeitung von ZUGFeRD-Dokumenten aktivieren.

ZUGFeRD-Verarbeitung

"*Individuelles Zielverzeichnis*"

Es kann ein vom Standard-Zielverzeichnis abweichendes Zielverzeichnis für die ZUGFeRD-Verarbeitung angegeben werden. Alle ZUGFeRD-Dokumente und Rechnungen werden in dieses Verzeichnis geschrieben.

"ZUGFeRD-Rechnungen extrahieren"

Sie können bei der ZUGFeRD-Verarbeitung entweder nur die erkannten Dokumente in das Zielverzeichnis verschieben oder über diese Option auch die eingebettete Rechnung im XML-Format extrahieren. Die Rechnung wird immer nach dem Dokumentennamen benannt.

Bei existierendem Dateinamen

Hier wird gesteuert was passieren soll, wenn im Zielverzeichnis bereits eine Datei oder eine XML-Rechnung gleichen Namens vorhanden ist.

"neue ins Fehlerverzeichnis"
Wenn der Zieldateiname existiert, wird das Dokument in das Fehlerverzeichnis geschrieben.

"fortlaufende Nummer vergeben
Bei der Option "fortlaufende Nummer vergeben" bilden sich die Dateinamen immer aus <DATEINAME>.NNNN.<DATEIENDUNG>. Bereits die erste Zieldatei erhält schon eine fortlaufende Nummer NNNN=0001 im Dateinamen.

"überschreiben"
Die im Zielverzeichnis existierenden Dateien werden überschrieben.

Dokumente ohne ZUGFeRD-Informationen

Hier wird gesteuert was passieren soll, wenn im Zielverzeichnis bereits eine Datei oder eine XML-Rechnung gleichen Namens vorhanden ist.

"neue ins Fehlerverzeichnis"
Wenn der Zieldateiname existiert, wird das Dokument in das Fehlerverzeichnis geschrieben.

"fortlaufende Nummer vergeben
Bei der Option "fortlaufende Nummer vergeben" bilden sich die Dateinamen immer aus <DATEINAME>.NNNN.<DATEIENDUNG>. Bereits die erste Zieldatei erhält schon eine fortlaufende Nummer NNNN=0001 im Dateinamen.

Dokumente ohne ZUGFeRD-Informationen

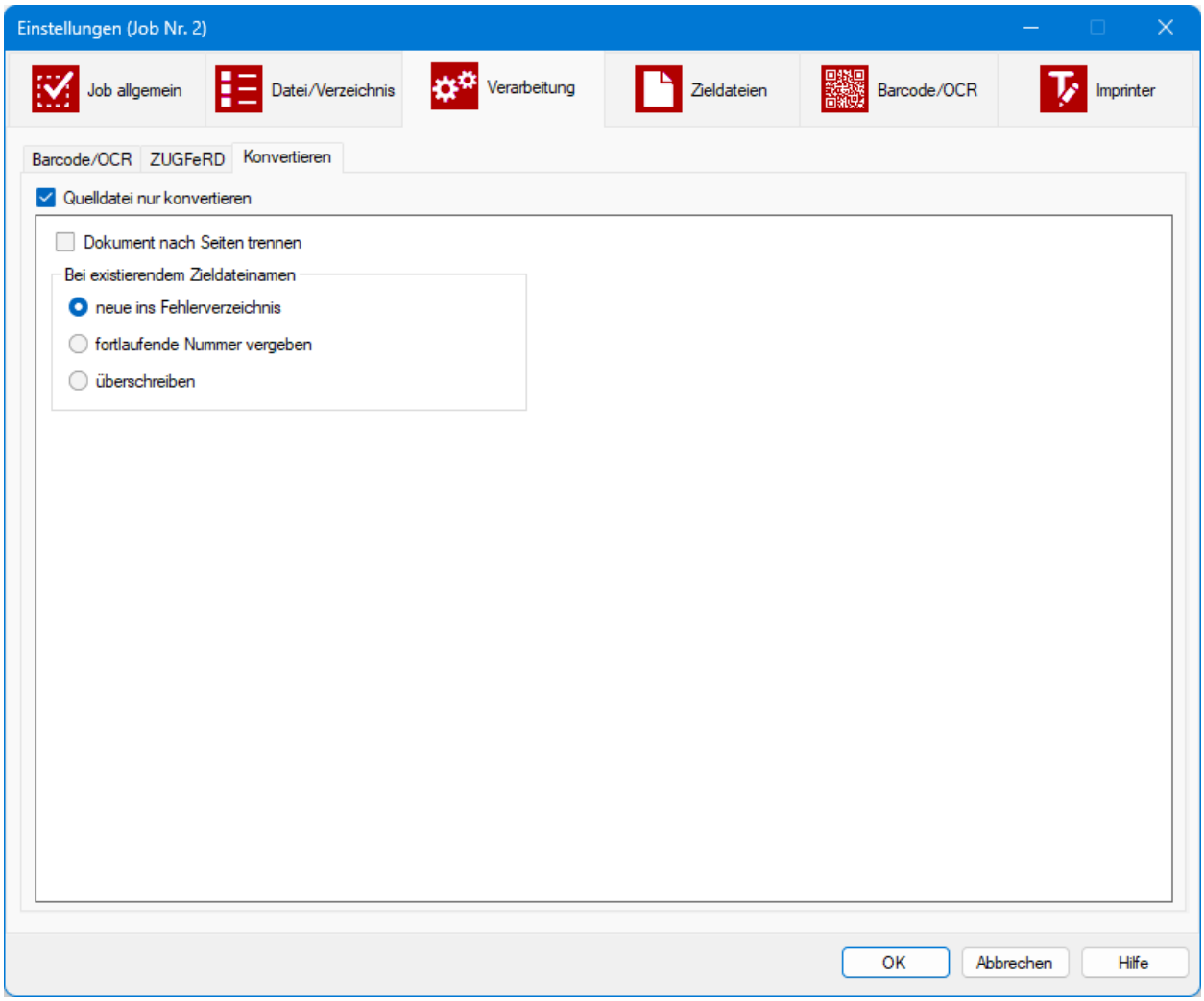
Hier wird gesteuert was passieren soll, wenn keine ZUGFeRD-Rechnung im Dokument gefunden wird.

"in das Fehlerverzeichnis verschieben"
Die Quelldatei wird in das Fehlerverzeichnis verschoben

"Barcode-/OCR-Verarbeitung"
Wenn keine ZUGFeRD-Daten gefunden werden, wird die Datei für die Barcode-/OCR-Verarbeitung im Quellverzeichnis belassen.

3.2.3.5 Konvertierung des Dateiformates

Der DocumentAssembler kann auch dazu eingesetzt werden, um Dateien von einem Format in ein anderes zu konvertieren. Eine Barcodeerkennung wird hierbei nicht durchgeführt. Dabei ist es möglich, mehrseitige Dokumente in einzelne Seiten aufzuteilen oder das gesamte Dokument zu konvertieren.



"Quelldatei nur konvertieren"

Die Quelldatei wird nur in das gewünschte Format unter Umgehung der Barcodeerkennung konvertiert.

"Dokument nach Seiten trennen"

Ein mehrseitiges Quelldokument wird in einzelne Seiten getrennt. Die Zieldateien bekommen den Namen der Quelldatei plus lfd. Nummer.

Bei existierendem Dateinamen

Hier wird gesteuert was passieren soll, wenn im Zielverzeichnis bereits eine Datei gleichen Namens vorhanden ist.

"neue ins Fehlerverzeichnis"

Wenn der Zieldateiname existiert, wird das Dokument in das Fehlerverzeichnis geschrieben.

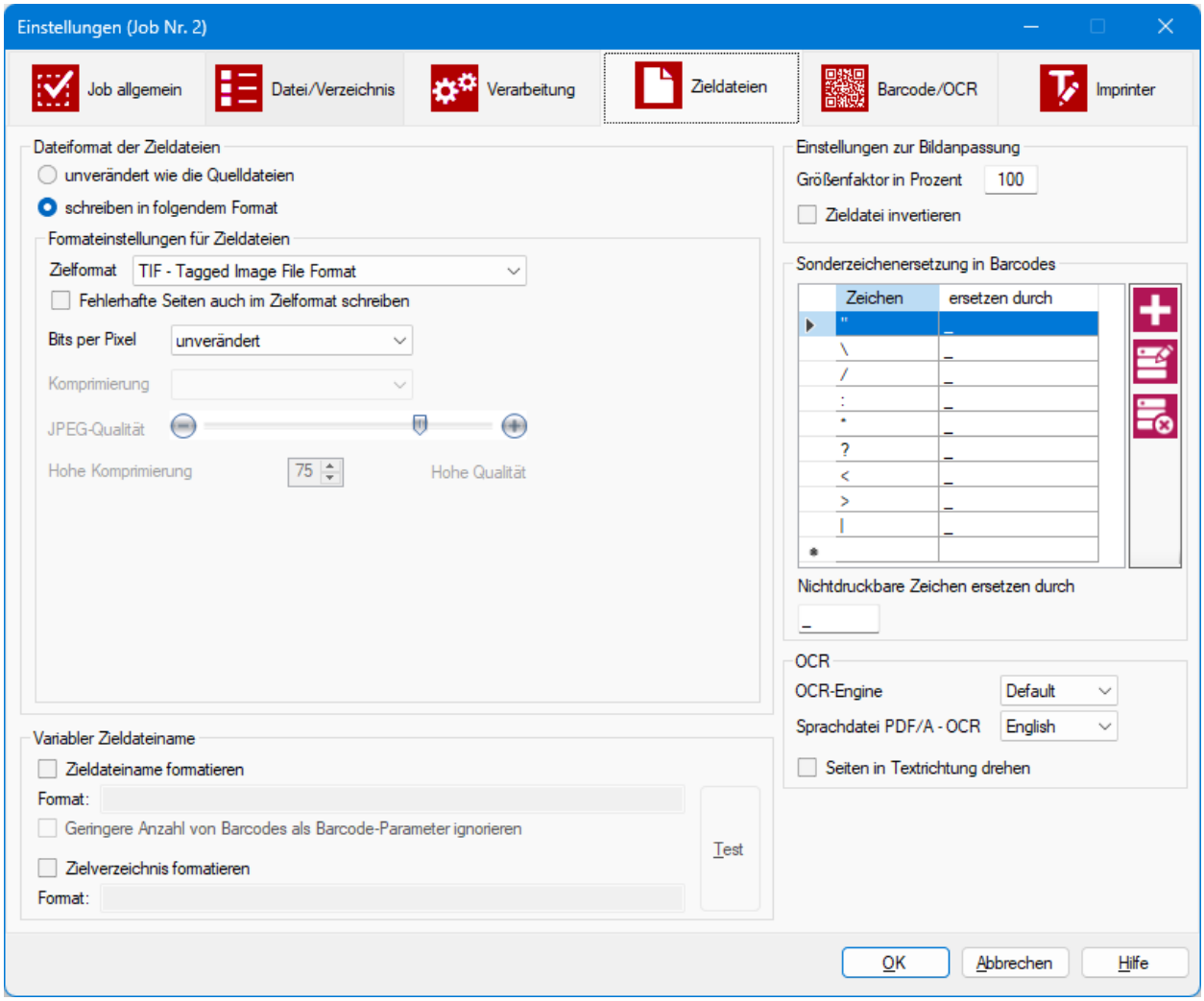
"fortlaufende Nummer vergeben

Bei der Option "fortlaufende Nummer vergeben" bilden sich die Dateinamen immer aus <DATEINAME>.NNNN.<DATEIENDUNG>. Bereits die erste Zieldatei erhält schon eine fortlaufende Nummer NNNN=0001 im Dateinamen.

"überschreiben"

Die im Zielverzeichnis existierende Dateie wird überschrieben.

3.2.4 Registerkarte "Zieldateien"



Auf dieser Registerkarte kann bestimmt werden, ob und wie die Zieldateien nach der Barcodeerkennung in ein anderes Dateiformat konvertiert werden sollen. Die Standardeinstellung ist, dass das Dateiformat der Quelldateien unverändert bleibt.

Der QS-DocumentAssembler unterstützt die Formate BMP, JPG, GIF, PNG und TIF als Zielformate für Rastergrafiken sowie das Adobe PDF-Format

Wann könnte die Formatkonvertierung nützlich sein?

- Wenn Ihre Quelldateien unterschiedliche Dateitypen haben, z.B. mit einem Dokumenten-Scanner erfasste TIFs und mit Digitalkameras erfasste JPGs, dann kann die Konvertierung dafür sorgen, dass die Zieldateien einheitlich sind. Für Folgeanwendungen, z.B. Archivsysteme, ist dies oft von Vorteil.
- Wenn Sie, um eine zuverlässige Barcodeerkennung zu gewährleisten, mit einer recht hohen Auflösung scannen, können Sie durch das Konvertieren nicht nur das Dateiformat ändern, sondern auch die Bildauflösung reduzieren.
- Sie können die Bilder nach dem Scannen zunächst mit einer geringen Kompressionsrate abzuspeichern. Dadurch wird der Informationsverlust geringgehalten und die Barcodeerkennung kann zuverlässig durchgeführt werden. Nach der Verarbeitung spei-

chern Sie dann die Bilder stark komprimiert, ohne Rücksicht auf die Auswirkungen, die dies auf die bereits erfolgte Barcodeerkennung hat.

Formateinstellungen für Zieldateien

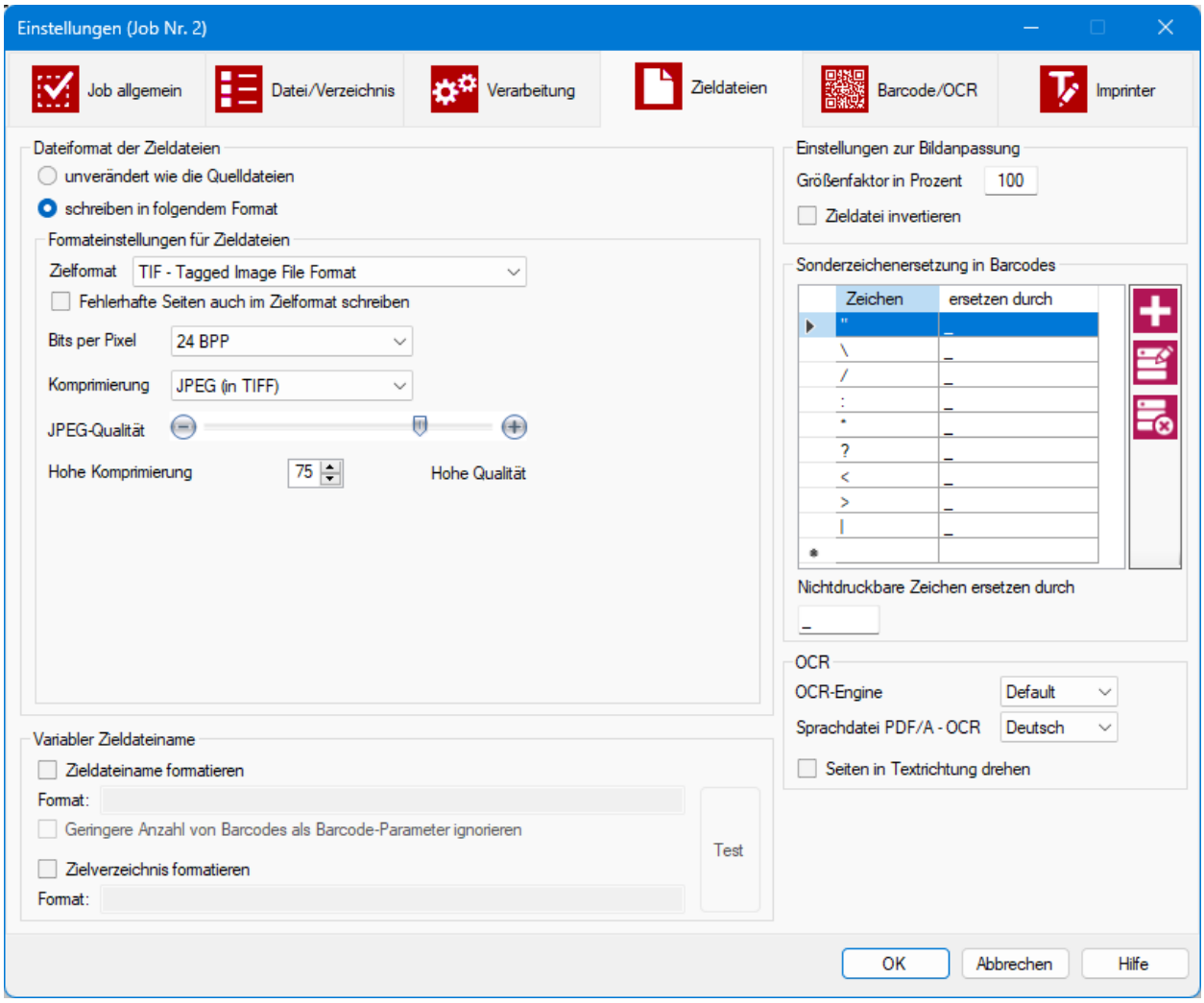
Je nach gewähltem Zielformat stehen einige Optionen, die auf dem Dialog zur Auswahl stehen, nicht zur Verfügung.

"schreiben in folgenden Format"
: Auswahl des gewünschten Zielformates, entweder ein genanntes Rasterformat oder Adobe PDF. Je nach gewähltem Format stehen einige der folgenden Optionen nicht zur Verfügung.

"Fehlerhafte Seiten auch im Zielformat schreiben."
: Fehlerhafte Seiten werden standardmäßig im Quellformat in das Fehlerverzeichnis geschrieben. Mit dieser Option haben Sie die Möglichkeit, diese Seiten auch in das definierte Zielformat zu wandeln.

Zielformat Rastergrafik (BMP, JPG, GIF, PNG, TIF)

Der QS-DocumentAssembler unterstützt die Raster-Formate BMP, JPG, GIF, PNG und TIF als Zielformate für Rastergrafiken, wobei nur das TIF-Format mehrseitige Zieldokumente unterstützt. Daher wird im Stapelmodus nur das TIF-Format als Rasterformat angezeigt.



"Bits Per Pixel"

Die Farbtiefe für die Zieldateien wird hier gewählt. Bei "*unverändert*" wird die Farbtiefe der Quelldatei übernommen, es sei denn diese Farbtiefe wird vom gewählten Format nicht unterstützt. Dann wird automatisch gewählt: Wenn möglich die nächsthöhere Farbtiefe, die unterstützt wird. Sonst die nächstniedrigere.

"*Komprimierung*"

Falls vorhanden können die Komprimierungs-Formate des Dateityps eingestellt werden.

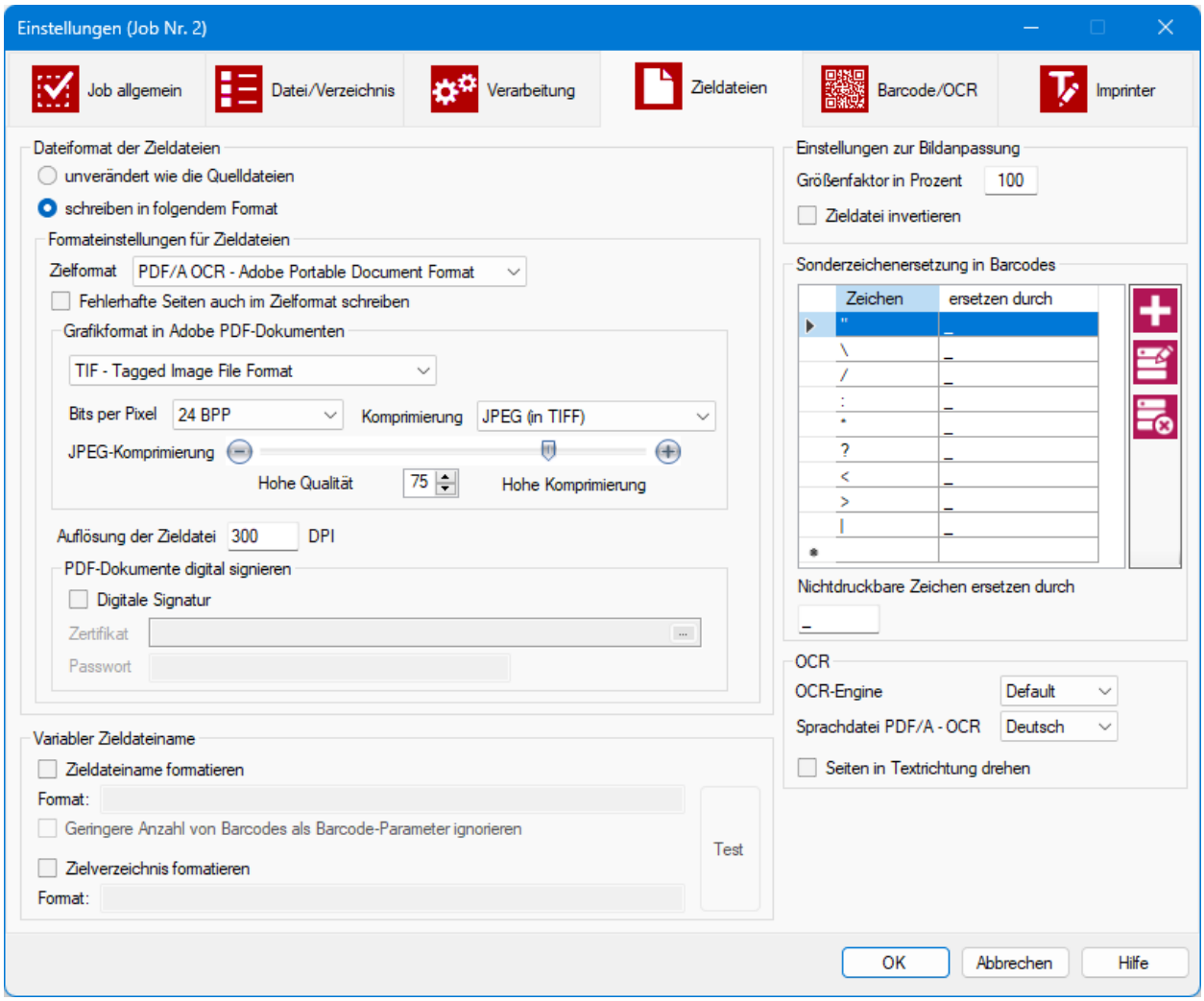
"JPEG-Qualität"

Bei JPEG-Komprimierung kann für JPG-Formate zwischen 0 (hohe Komprimierung) und 100 (geringe Komprimierung) gewählt werden.

Zielformat PDF / PDF/A / PDF/A - OCR

Bei der expliziten Umwandlung in ein PDF-Dokument können auch PDF/A-Dokumente erzeugt werden. Bei der Umwandlung in ein PDF/A-Dokument haben Sie die Möglichkeit, den Text des Dokumentes über einen automatischen OCR-Prozess im PDF-Dokument zu speichern bzw. das Dokument ohne OCR als PDF/A-Dokument auszugeben.

Bitte beachten Sie, dass der OCR-Prozess sehr zeitaufwendig ist und die Verarbeitungsdauer sich hierdurch wesentlich erhöhen kann.



Formateinstellungen für Zieldateien

"Grafikformat in Adobe PDF-Dokumenten"

Sie können aus einer Liste von gängigen Raster-Dateiformaten auswählen, welches Format im PDF-Dokument eingebettet wird. Je nach gewähltem Format stehen folgende Parameter zur Verfügung.

"Bits Per Pixel"

Die Farbtiefe für die Zieldateien wird hier gewählt. Bei "*unverändert*" wird die Farbtiefe der Quelldatei genutzt, es sei denn diese Farbtiefe wird vom gewählten Format nicht unterstützt. Dann wird automatisch gewählt: Wenn möglich die nächsthöhere Farbtiefe, die unterstützt wird. Sonst die Nächstniedrigere.

"Komprimierung"

Falls vorhanden können die Komprimierungs-Formate des Dateityps eingestellt werden.

"JPEG-Qualität

Bei JPEG-Komprimierung kann für JPG-Formate zwischen 0 (hohe Komprimierung) und 100 (geringe Komprimierung) gewählt werden.

"Auflösung der Zieldatei

Geben Sie an, mit welcher Auflösung die Zieldatei geschrieben werden soll. Bei einem von den Quell-DPI abweichenden Wert wird die Grafik entsprechend skaliert.

PDF-Dokumente digital signieren

Sie können PDF-Dokumente während der Verarbeitung mit eigenen digitalen Zertifikaten signieren. Ihre Dokumente können anschließend noch verändert werden. Die Signatur wird dadurch ungültig. Dokumente können mehrfach signiert werden. Die Veränderungen an den Dateien können so abgesichert und erkannt werden.

Dem DocumentAssembler liegt ein Test-Zertifikat bei, das Sie für Ihren Test benutzen können. Sie finden dieses Test-Zertifikat im Verzeichnis "\Certificate"

"Zertifikat"
: Wählen Sie hier ihr Zertifikat aus.

"Passwort"
: Geben Sie hier ihr Passwort für das Zertifikat ein. Das Passwort wird durch den DocumentAssembler verschlüsselt gespeichert.

OCR

Bei der Erstellung von durchsuchbaren PDF-Dokumenten können Sie zwischen der Tesseract-Engine und unserer Standard-Engine wählen.

"OCR-Engine"
: Sie können wählen zwischen der schnelleren Tesseract-Engine und der sehr viel genaueren Standard-Engine.

"Sprachdatei PDF/A – OCR"
: Bei der Texterkennung von gescannten Dokumenten gewährleistet eine Sprachdatei, dass die länderspezifischen Besonderheiten des Textes erhalten bleiben.

"Seiten in Textrichtung drehen"
: Wenn Seiten z.B. verkehrt herum gescannt werden, werden die Dokumente in Textrichtung gedreht.

Einstellungen zur Bildanpassung

"Größenfaktor in Prozent"
: Bei einem Wert ungleich 100 wird die Zieldatei entsprechend vergrößert bzw. verkleinert.

"Zieldatei invertieren"
: Die Datei wird invertiert abgelegt.

Variabler Zieldateiname

Wenn Sie im Dateinamen abweichend vom Standard zusätzliche Informationen unterbringen wollen, haben Sie eine Reihe von Parametern zur Auswahl:

%QUELLDATEI%	Ursprünglicher Dateiname
%BARCODE%	Der erkannte erste Barcode
%BARCODEn%	Der erkannte n-te Barcode
%BC_REGEX%	Ergebnis eines Regulären Ausdrucks auf den ersten Barcode angewandt.
%BC_REGEX_REPLACE%	Ersetzung durch einen Regulären Ausdrucks im ersten Barcode

	ersetzt.
%SEITEN_NR%	Die aktuelle Seitennummer des Quelldokumentes
%SEITEN_NRn%	Die aktuelle Seitennummer des Quelldokumentes, formatiert auf n Stellen, z.B. %SEITEN_NR3% entspricht "001"
%SEITENZAHL%	Die Seitenzahl des Zieldokumentes
%SEITENZAHLn%	Die Seitenzahl des Zieldokumentes, formatiert auf n Stellen, z.B. %SEITENZAHL3% entspricht "005"
%LFDNR%	Anzahl der Vorkommen eines Dateinamens im Zielverzeichnis
%LFDNRn%	Die lfd. Nummer formatiert auf n Stellen, z.B. %LFDNR4% entspricht "0001"
%YYYY%	Jahr vierstellig. z.B. 2010
%YY%	Jahr zweistellig, z.B. 10
%Y%	Tag im Jahr
%MMMM%	Monat Langform, z.B. "September"
%MMM%	Monat als Kürzel, z.B. "Sep"
%MM%	Monat als Ziffer mit führender Null, z.B. "07"
%M%	Monat einstellig
%DDDD%	Tag Langform, z.B. "Dienstag"
%DDD%	Tag als Kürzel, z.B. "Di"
%DD%	Tag als Ziffer mit führender Null, z.B. "01"
%D%	Tag, einstellig
%TDJ%	Tag des Jahres, z.B. 125
%HH%	Stunde, zweistellig
%H%	Stunde, einstellig
%mm%	Minute, zweistellig
%m%	Minute, einstellig
%SS%	Sekunde, zweistellig
%S%	Sekunde, einstellig
%AM/PM	Zeitangabe AM/PM, z.B. "AM"
%DATUM_KURZ%	Das aktuelle Datum im Format TT.MM.JJJJ
%DATUM_LANG%	Datum in Langformat, z.B. Dienstag, 26. November 2013
%ZEIT%	Die aktuelle Uhrzeit im Format HHmmSS
%ZEIT%	Die aktuelle Uhrzeit im Format HHmmSS
%ZEIT%	Die aktuelle Uhrzeit im Format HHmmSS

Im Zieldateinamen muss mindestens eine der folgenden Variablen vorhanden sein:"%QUELLDATEI%", "%BARCODE%", "%BC_REGEX%", "%BC_REGEX_REPLACE%".
Wenn mehrere auf der Seite vorkommende Barcodes zur Bildung des variablen Zieldateinamens herangezogen werden sollen, können Sie die Barcodes über den Parameter %BARCODEn%, z.B. %BARCODE1%, %BARCODE2%, etc. auswählen.
Die Barcodes werden in der Reihenfolge übernommen, in der Sie im Dokument gefunden werden. Diese muss nicht notwendigerweise mit der Reihenfolge der Barcodes im Dokument übereinstimmen. Sie sollten daher die Barcodes auf der Seite anhand ihrer Position sortieren lassen (siehe Kapitel 3.2.5, Registerkarte "Barcode/OCR").
Zusätzlich können Sie beliebigen Text in den Dateinamen aufnehmen, solange die Zeichen für die Erstellung von Dateinamen zugelassen sind.

Bitte beachten Sie, dass nicht alle Parameter (insbesondere zeitbasierte Parameter wie Sekunden im Stapelmodus) bei allen Optionen der Verarbeitungsmodi sinnvoll sind.

"Geringere Anzahl von Barcodes als Barcode-Parameter ignorieren "

Wenn weniger Barcodes gefunden werden als Barcode-Parameter im variablen Zieldateinamen angegeben sind, wird die Datei in das Fehlerverzeichnis verschoben, da der Zieldateiname nicht komplett aufgebaut werden kann. Über diese Option können Sie den Dateinamen auch erzeugen, wenn weniger Barcodes gefunden werden, als Barcode-Parameter vorhanden sind. Die Überzähligen Barcode-Parameter werden dann aus den Zieldateinamen entfernt.

Variables Zielverzeichnis

Das Zielverzeichnis kann variabel mit allen o.g. Parameter definiert werden. Der angegebene variable Teil wird dabei an das unter "Pfade und Dateinamen" definierte Zielverzeichnis angehängt.

Variabler Barcode

Aus dem Barcode kann über verschiedene Parameter ein bestimmter Bereich für den variablen Zieldateinamen, für das variable Zielverzeichnis sowie für den Imprinter ausgeschnitten werden.

BC_SUBSTR	Teilzeichenfolge ab einer bestimmten Position mit einer bestimmten Länge.
BC_STARTSWITH	Teilzeichenfolge ab einem bestimmten Teilstring mit einer bestimmten Länge
BC_ENDSWITH	Teilzeichenfolge vor einem bestimmten Teilstring mit einer bestimmten Länge
BC_BETWEEN	Teilzeichenfolge zwischen zwei definierten Teilstrings
BC_REPLACE	Teilzeichenfolge aus dem Barcode ersetzen.
BC_REGEX	Ergebnis eines Regulären Ausdrucks auf den ersten Barcode angewandt.
BC_REGEX_REPLACE	Ersetzung des Ergebnisses eines Regulären Ausdrucks im ersten Barcode durch einen String.

Beispiele:

%BC_SUBSTR(5)%	Ab Position 5 des Barcodes bis zum Ende
0123456789012	
%BC_SUBSTR(-5)%	Alle Zeichen ab dem fünftletzten Zeichen bis zum Anfang
0123456789012	
%BC_SUBSTR(0,5)%	5 Zeichen ab Position 0 (Anfang) des Barcodes

0123456789012	
%BC_SUBSTR(10,-5)%	5 Zeichen rückwärts ab Position 10
0123456789012	
%BC_SUBSTR(-1,-5)%	5 Zeichen rückwärts ab letztem Zeichen (-1)
0123456789012	
%BC_SUBSTR(-5,4)%	4 Zeichen ab fünftletzter Position
0123456789012	
%BC_STARTSWITH(XXX)%	Alle Zeichen ab dem Vorkommen von "XXX" bis zum Ende
012XXX3456789012	
%BC_STARTSWITH(XXX,5)%	5 Zeichen ab dem Vorkommen von "XXX"
012XXX3456789012	
%BC_ENDSWITH(XXX)%	Alle Zeichen vor dem ersten Auftreten von "XXX" bis zum Anfang
012345678XXX9012	
%BC_ENDSWITH(XXX,5)%	5 Zeichen rückwärts vom ersten Auftreten von XXX
012345678XXX9012	
%BC_BETWEEN(XXX,YYY)%	Alle Zeichen zwischen "XXX" und "YYY"
012XXX34567890YYY12	
%BC_REPLACE(XXX,YYY)%	Ersetze "XXX" durch "YYY"
012XXX3456789012 => 012YYY3456789012	
%BC_REPLACE(3,5,YYY)%	Ersetze ab Position 3 in der Länge 5 durch "YYY".
012XXX3456789012 => 012YYY56789012	
%BC_REGEX(ABCD\d*(?=_*))%	Verarbeitung von strukturell ähnlichen Barcodes mit unterschiedlicher Länge. Alle Barcodes, die mit "ABCD" anfangen und eine beliebige Anzahl an Ziffern haben bis zum Unterstrich. Ist kein Unterstrich vorhanden, wird bis zum Ende gelesen.
ABCD12345678_99 => ABCD12345678 ABCD987654321 => ABCD987654321	
%BC_REGEX_REPLACE(AA[3,4],XX1)%	Ersetzen von über einen Regulären Ausdruck gefunden Zeichenketten mit eigener Zeichenkette. Gesucht wird nach den Zeichenketten AA3/AA4, welche durch XX1 ersetzt werden.
AA3BB6CC7 => XX1BB6CC7 AA4BB5CC6 => XX1BB5CC6	

Variabler Quelldateiname

Auch der Quelldateiname kann entsprechend mit QD_SUBSTR() behandelt werden. Das Vorgehen ist analog zur oben dargestellten Vorgehensweise.

Bei der Benutzung der Parameter gilt es grundsätzlich zu beachten, dass der Barcode bzw. der Quelldateiname auch die im Parameter benutzen Längen aufweist. Ist dies nicht der Fall wird ein Fehler protokolliert und die Verarbeitung wird angehalten.

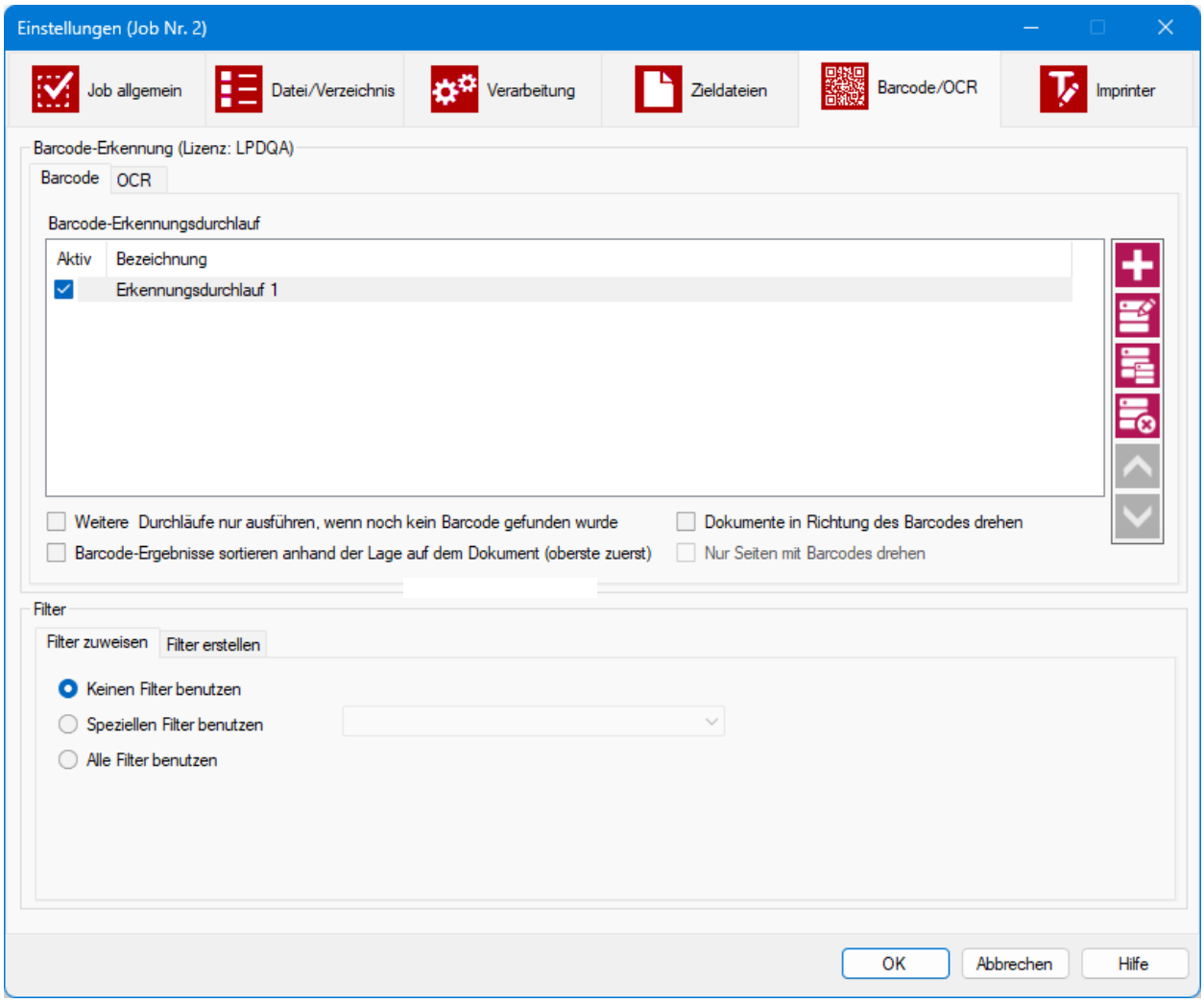
Sonderzeichenersetzung in Barcodes

Wenn sich der Dateiname aus dem erkannten Barcode ergeben soll, dürfen im Barcode keine Sonderzeichen enthalten sein, sonst kann die Datei nicht umbenannt werden. Um dieses Problem bei der Umbenennung der Datei zu umgehen, kann man mögliche in Barcodes vorkommende Sonderzeichen durch selbst definierte Zeichen ersetzen.

"Nichtdruckbare Zeichen ersetzen durch"

Für die Umbenennung nach Barcode werden die ASCII- Werte 00 bis 31 durch ein frei definierbares Ersatzzeichen ersetzt.

3.2.5 Registerkarte "Barcode/OCR"



Der Überschrift können Sie entnehmen, welche Barcode-Lizenz auf dem System vorhanden ist. Der Demo-Modus wird zusätzlich durch einen roten Schriftzug "DEMO" im linken unteren Bereich des Dialoges gekennzeichnet, wenn eine DEMO-Lizenz gefunden wurde.

Der DocumentAssembler unterstützt neben der Erkennung von Barcodes auf Dokumenten auch Erkennungsdurchläufe zur Erkennung von Textinhalten (OCR) auf Dokumenten, deren Ergebnisse wie die Inhalte von Barcodes für die weitere Verarbeitung der Dokumente verwendet werden können.

Die Konfiguration erfolgt jeweils in separaten Barcode- und OCR-Erkennungsdurchläufen.

3.2.5.1 Barcode-Erkennung

Auf dem Reiter "Barcode" konfigurieren Sie die Barcode-Erkennung über Erkennungsdurchläufe. Ein Erkennungsdurchlauf beinhaltet die für die Barcodesuche relevanten Parameter, mit denen die Barcodesuche durchgeführt wird.

In den meisten Fällen reicht ein einzelner Erkennungsdurchlauf aus. Gelegentlich kommt es vor, dass die Qualität der Quelldateien sehr unterschiedlich ist, so dass die Erkennung mit unterschiedlichen Parametern durchgeführt werden muss. Wenn Sie diese unterschiedlichen

Parameter nicht in einem Erkennungsdurchlauf abbilden können, erstellen Sie einfach einen zusätzlichen Erkennungsdurchlauf.

Wenn Sie auf den Dokumenten sowohl lineare als auch 2D Barcodes suchen wollen oder wenn Sie zwei unterschiedliche 2D Barcode Typen erkennen wollen, müssen Sie dies über zwei Erkennungsdurchläufe konfigurieren.

Unterschiedliche lineare Barcode Typen können problemlos in einem einzigen Erkennungsdurchlauf gelesen werden. Wenn Sie mehrere Erkennungsdurchläufe erstellt haben, für bestimmte Suchvorgänge aber nicht alle Durchläufe benötigen, können Sie einzelne Erkennungsdurchläufe über die Checkbox vor dem Namen des Durchlaufes deaktivieren.

Die Bearbeitung der Erkennungsdurchläufe erfolgt über die vertikale Buttonleiste rechts neben der Liste der Erkennungsdurchläufe.

Ein neuer Erkennungsdurchlauf wird erstellt. Der Erkennungsdurchlauf wird mit Standard-Parametern vorbelegt und der Dialog für die Barcode-Einstellungen wird geöffnet.

Ruft den Dialog für die Bearbeitung des markierten Erkennungsdurchlaufs auf.

Duplizieren des aktuell markierten Erkennungsdurchlaufes.

Mit dieser Schaltfläche löschen Sie den in der Liste markierten Erkennungsdurchlauf.

Mit diesen Pfeil-Tasten können Sie die Ausführungsreihenfolge der Erkennungsdurchläufe anpassen. Markieren Sie einen Eintrag in der Liste, um ihn anschließend durch Anwahl der Pfeil-Tasten zu verschieben.

"Weitere Durchläufe nur ausführen, wenn noch kein Barcode gefunden wurde"

Wenn Sie diese Option aktivieren, wird die Barcodeerkennung gestoppt, sobald nach einem Erkennungsdurchlauf mindestens ein Barcode-Ergebnis für die aktuelle Seite vorliegt. Die übrigen Erkennungsdurchläufe werden dann für diese Seite nicht mehr ausgeführt. Ist diese Option nicht aktiviert, werden immer alle Erkennungsdurchläufe durchgeführt.

"Barcode-Ergebnisse sortieren anhand der Lage auf dem Dokument"

Ohne Aktivierung dieser Option werden die erkannten Barcodes in der Reihenfolge berichtet, in der Sie gefunden wurden. Die Reihenfolge bestimmt sich so durch den Erkennungsdurchlauf (alle Ergebnisse des 1. Durchlaufs, dann alle des 2. Durchlaufs usw.), innerhalb eines Durchlaufs ist die Reihenfolge willkürlich (meist von oben nach unten, das muss aber nicht so sein).
Ist die Option markiert, wird nach der Durchführung aller Erkennungsdurchläufe eine Sortierung anhand der Barcode-Positionen durchgeführt. Der oberste Barcode auf der Seite wird zuerst berichtet. Dabei erfolgt *keine* Analyse, ob die Seite evtl. gedreht ist. Oben entspricht dem, was bei der Darstellung der Seite auf dem Bildschirm oben ist.

"Dokumente in Richtung des Barcodes drehen"

Falls Dokumente versehentlich auf dem Kopf stehend gescannt wurden, kann das Dokument über die Barcodeausrichtung gedreht werden.

"Nur Seiten mit Barcodes drehen"

Über diese Option können Sie einschränken, dass nur Seiten, auf denen sich ein Barcode befindet, gedreht werden. Seiten ohne Barcode sind davon nicht betroffen.

3.2.5.2 OCR-Erkennung

Auf dem Reiter "OCR" konfigurieren Sie die OCR-Erkennung über Erkennungsdurchläufe. Ein Erkennungsdurchlauf beinhaltet die für die OCR-Suche relevanten Parameter, mit denen die OCR-Suche durchgeführt wird.

Die Bearbeitung der Erkennungsdurchläufe erfolgt über die vertikale Buttonleiste rechts neben der Liste der Erkennungsdurchläufe.

Ein neuer Erkennungsdurchlauf wird erstellt. Der Erkennungsdurchlauf wird mit Standard-Parametern vorbelegt und der Dialog für die OCR-Einstellungen wird geöffnet.

Ruft den Dialog für die Bearbeitung des markierten Erkennungsdurchlaufs auf.

Duplizieren des aktuell markierten Erkennungsdurchlaufes.

Mit dieser Schaltfläche löschen Sie den in der Liste markierten Erkennungsdurchlauf.

Mit diesen Pfeil-Schaltflächen können Sie die Ausführungsreihenfolge der Erkennungsdurchläufe anpassen. Markieren Sie einen Eintrag in der Liste, um ihn anschließend durch Anwahl der Pfeil-Tasten zu verschieben.

"Weitere Durchläufe nur ausführen, wenn noch kein Text gefunden wurde"

Wenn Sie diese Option aktivieren, wird die OCR-Erkennung gestoppt, sobald nach einem Erkennungsdurchlauf mindestens ein Ergebnis für die aktuelle Seite vorliegt. Die übrigen Erkennungsdurchläufe werden dann für diese Seite nicht mehr ausgeführt. Ist diese Option nicht aktiviert, werden immer alle Erkennungsdurchläufe durchgeführt.

"Dokumente in Richtung des Textes drehen"

Falls Dokumente versehentlich auf dem Kopf stehend gescannt wurden, kann das Dokument über die Textausrichtung gedreht werden.

"Nur Seiten mit Treffern drehen"

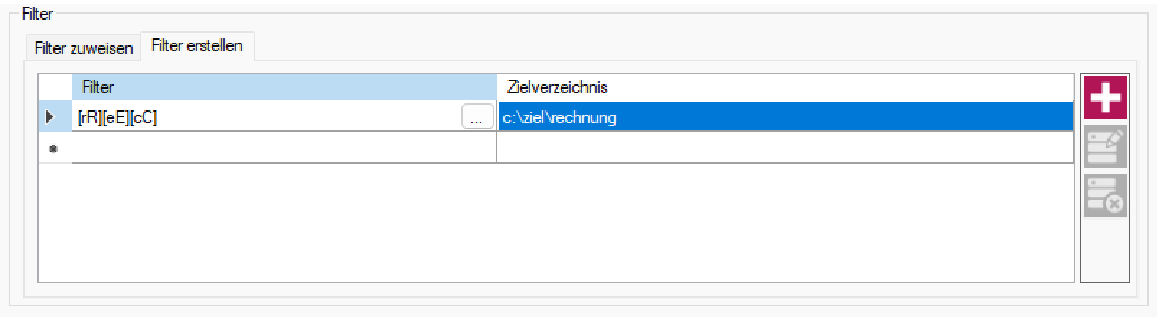
Über diese Option können Sie einschränken, dass nur Seiten, auf denen Text erkannt wurde, gedreht werden.

3.2.5.3 Filter

Es kann bei Dokumenten vorkommen, dass sie Barcodes aus mehreren Quellen enthalten. Mit der Filterfunktion kann die Erkennung auf die benötigten Barcodes eingeschränkt werden. Auch die aus der OCR-Erkennung gewonnen Ergebnisse können über die Filter noch nachträglich gefiltert werden bzw. zum Verschieben in filterspezifische Verzeichnisse benutzt werden.

Erkannte Barcodes und Ergebnisse aus der OCR-Erkennung werden vom QS-DocumentAssembler vor jeder weiteren Verarbeitung einer Filterung unterzogen. Entspricht ein Barcode/Ergebnis nur einem von eventuell mehreren definierten Filtern, so wird er der Liste der erkannten Barcodes/Ergebnisse hinzugefügt. Filter werden intern also mit "oder" verknüpft. Alle anderen Barcodes/Ergebnisse werden in der weiteren Verarbeitung nicht berücksichtigt.

Registerkarte "Filter erstellen"

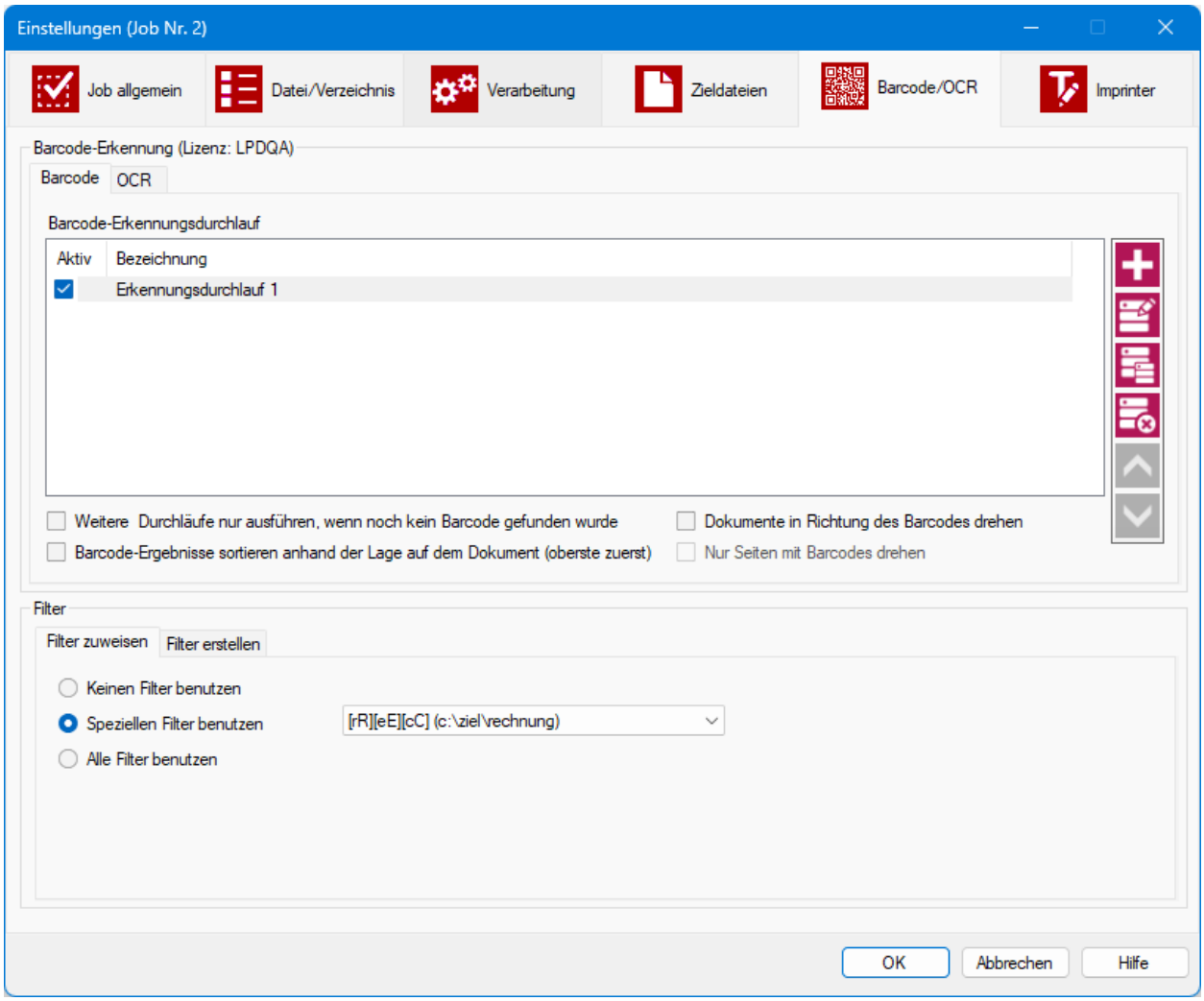


Ein- oder mehrere Barcode-Filter können über den Reiter "Filter erstellen" hinzugefügt werden. Dabei kann jedem Filter ein individuelles Zielverzeichnis für die Erstellung des Zieldokumentes zugeordnet werden. Dieses individuelle Zielverzeichnis überschreibt dann das auf der Registerkarte "Datei/Verzeichnis" definierte Zielverzeichnis.

Die Eingabe des Filters sowie des Zielverzeichnisses erfolgt durch direkte Dateneingabe in der Liste der Filter. Zusätzlich kann über einen Button am rechten Rand des jeweiligen Eingabefeldes ein Fenster für die Bearbeitung eines längeren Regulären Ausdrucks für den Filter bzw. ein Auswahldialog für das Verzeichnis geöffnet werden.

Über die Button-Leiste am rechten Rand der Filterliste kann der ausgewählte Filter gelöscht oder auch bearbeitet werden.

Es wird empfohlen, zunächst durch die Wahl der Barcodeparameter (Typ, Länge usw.) /OCR-Parameter die Menge der erkennbaren Barcodes/Ergebnisse zu reduzieren und sie erst dann durch Filter weiter einzuschränken. Die Filter müssen auf dem Reiter "Filter zuweisen" den einzelnen Erkennungsdurchläufen zugeordnet werden.



Filter zuweisen

Es reicht nicht aus, die Filter nur zu erstellen. Um die Filter zu nutzen, müssen Sie einem Erkennungslauf zugeordnet werden. Wählen Sie hierzu bitte in der Liste der Erkennungsdurchläufe den gewünschten Durchlauf aus, dem Sie Filter zuweisen wollen. Sie haben dabei die Möglichkeit, dem Erkennungsdurchlauf einen speziellen Filter, den Sie aus der entsprechenden Combobox auswählen, zuzuordnen oder dem Erkennungsdurchlauf alle vorhandenen Filter zuzuordnen.

"Keinen Filter benutzen"
: Bei der Barcodesuche werden keine Filter benutzt.

"Speziellen Barcodefilter benutzen"
: Wenn Sie Barcodefilter erstellt haben, können Sie in der Combobox den gewünschten Filter zuweisen.

"Alle *Barcodefilter benutzen"*
: Alle definierten Filter werden benutzt.

Die Benutzung Regulärer Ausdrücke

Die Filter im DocumentAssembler werden über Reguläre Ausdrücke erstellt, die auf dem Regelwerk des Microsoft .NET Framework basieren. Weiterführende Informationen über die

Sprachelemente für Reguläre Ausdrücke im Microsoft .NET Framework finden Sie unter https://docs.microsoft.com/de-de/dotnet/standard/base-types/regular-expression-language-quick-reference

Die definierten Filter unterstützen einen Mustervergleich, wobei Platzhalterzeichen zugelassen sind. Dieser Mustervergleich bietet ein sehr breites Spektrum an Möglichkeiten.

Die folgende Tabelle zeigt einige ausgewählte Platzhalterzeichen, die Sie verwenden können, sowie die Anzahl der Ziffern oder Zeichenfolgen, für die eine Entsprechung erkannt wird.

Bei der Benutzung von Zeichen wird zwischen Groß- und Kleinschreibung unterschieden.

Zeichen im Muster	**Erkannte Entsprechung mit Ausdruck**
.	Beliebiges einzelnes Zeichen
*	Der voranstehende Ausdruck darf keinmal oder beliebig oft vorkommen.
+	Der voranstehende Ausdruck muss mindestens einmal vorkommen oder kann beliebig oft vorkommen.
\d	Beliebige einzelne Ziffer (0-9)
\D	Ein Zeichen, welches keine Ziffer ist
[Zeichenliste]	Ein oder mehrere beliebige Zeichen in Zeichenliste
[^Zeichenliste]	Ein oder mehrere beliebige Zeichen, die nicht in Zeichenliste enthalten sind
?	Zeichen davor kann vorkommen, muss es aber nicht
^	Beginn einer Zeichenkette, String muss mit folgendem String beginnen
$	Ende einer Zeichenkette, String muss mit vorherigem Zeichen/vorheriger Zeichenkette enden.
(?i)	Ignoriere Groß- und Kleinschreibung bei allen folgenden Zeichen.
(?-i)	Groß- und Kleinschreibung wird ab hier nicht mehr ignoriert.
|	Verknüpfung mehrerer Ausdrücke mit "oder"

Eine Gruppe aus einem oder mehreren Zeichen (Zeichenliste) in eckigen Klammern ([aei]) kann verwendet werden, um eine Entsprechung mit jedem einzelnen Zeichen im Ausdruck zu erkennen, wobei nur ein Zeichen aus der Zeichenliste treffen muss. Wenn Sie nach einer festen Zeichenkette suchen wollen, geben Sie die Zeichen ohne eckige Klammern ein.

Neben einer einfachen Liste aus Zeichen, die in eckige Klammern eingeschlossen ist, kann die Zeichenliste einen Bereich von Zeichen mit einem Bindestrich angeben([a-i]), der den obersten und den untersten Wert des Bereichs voneinander trennt. Bei der Verwendung von [A-Z] in Muster wird z.B. eine Entsprechung erkannt, wenn die entsprechende Zeichenposition in Ausdruck einen großgeschriebenen Buchstaben im Bereich von A bis Z enthält. Mehrere Bereiche können innerhalb der Klammern ohne Trennzeichen angegeben werden. Bei [a-zA-Z0-9] werden z.B. Entsprechungen (für ein Zeichen) mit allen alphanumerischen Zeichen erkannt.

Es gelten einige weitere wichtige Regeln für den Mustervergleich:

- Ein Negationszeichen (^) am Anfang von Zeichenliste ([^A-Z]) bedeutet, dass eine Entsprechung erkannt wird, wenn ein Zeichen in Ausdruck gefunden wird, das nicht in der Zeichenliste enthalten ist, hier also keine Großbuchstaben.
- Der Bindestrich (-) wird verwendet, um einen Bereich von Zeichen zu kennzeichnen.
- Wenn ein Zeichenbereich angegeben wird, müssen die Zeichen aufsteigend sortiert (A-Z oder 0-100) angegeben werden. [A-Z] ist ein gültiges Muster, aber [Z-A] nicht.

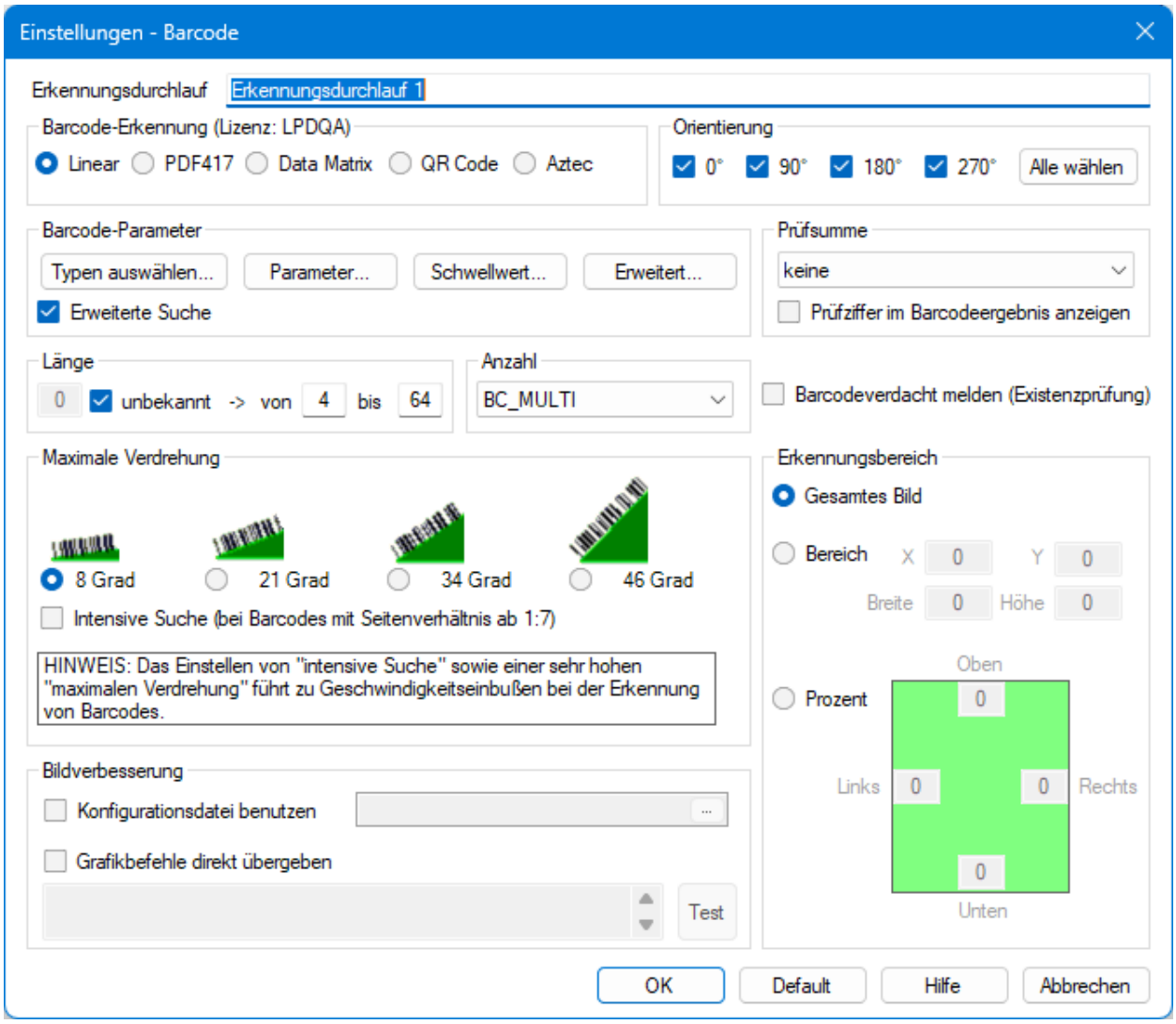
Möchte man Zeichen mit spezieller Bedeutung, sogenannte Metazeichen (z.B. Punkt, Bindestrich, Fragezeichen, etc.) als Ihren eigenen Wert benutzen, so muss man ihnen einen Backslash voranstellen. Dieses Voranstellen eines Escape-Zeichens zum Aufheben der Metabedeutung nennt man "escapen". Um z.B. nach einem Punkt zu suchen, benutzen Sie "\.". Wenn Sie nach einem Backslash suchen, so müssen Sie diesem auch einen Backslash voranstellen "\\".

Beispiele:

Filter	**Übereinstimmung**	**Erläuterung**
^RE	RE0001/REC	Am Anfang muss die Buchstabenkombination "RE" stehen.
RE$	0001RE	Die Buchstabenkombination "RE" muss am Ende stehen.
RE	RE001/ABRE001/001RE	Finden der Buchstabenkombination "RE" an beliebiger Stelle.
[rR][eE]	re01/AbRe01/01rE	Die Buchstabenkombination darf zusätzlich Kleinbuchstaben enthalten.
.\d\d\d	A123/0123/0x0123/AA123x	Ein beliebiges Zeichen gefolgt von drei Ziffern.
\d[AB]	1A4711/2B/012B	Beliebige Ziffer gefolgt vom Buchstaben "A" oder "B"
^1234$	1234	Exakte Übereinstimmung von Anfang bis Ende
^[0-8]*$	12345678/456	Nur Zahlen von 0-8 sind in beliebiger Häufigkeit zugelassen
^[^A-Z]*$	abc0123	Alle Zeichen außer Großbuchstaben sind zugelassen bei beliebiger Länge.
\D*\.\d*	Ein Test.000xx	Beliebige Anzahl nicht numerischer Zeichen, ein Punkt gefolgt von einer beliebigen Anzahl Ziffern.
REC(350|380)	REC350	REC350 oder REC380

3.2.5.4 Barcodeeinstellungen

Um zu den Barcodeeinstellungen zu gelangen müssen Sie entweder einen neuen Erkennungsdurchlauf anlegen oder einen bestehenden Erkennungsdurchlauf kopieren oder bearbeiten.



In diesem Einstellungsdialog legen Sie alle Details zur Barcodeerkennung im Erkennungsdurchlauf fest. Dazu zählen die Bezeichnung des Durchlaufs sowie alle Steuerparameter für die Barcode-Erkennung.

Die meisten Einstellungen betreffen die herkömmlichen linearen Barcodes (Strichcodes). Bei den 2D Barcodes (PDF417, Data Matrix, QR Code oder Aztec Code) werden fast alle Einstellungen automatisch definiert.

"Erkennungsdurchlauf"

Geben Sie hier einen beschreibenden Namen für den Erkennungsdurchlauf an.

"Barcode-Erkennung"

Je nach erworbener Programmlizenz können Sie hier zwischen der Suche nach linearen Barcodes, PDF 417, Data Matrix, QR Code oder Aztec Code wählen. Die verschiedenen Barcodetypen können nicht gleichzeitig in einem Durchlauf erkannt werden. Für jeden Barcodetyp müssen Sie einen separaten Erkennungsdurchlauf anlegen. Ver-

schiedene lineare Barcodetypen können hingegen in einem Erkennungsdurchlauf kombiniert werden. Ausnahmen hiervon sind der Pharma Code, Patch Code und der Codablock F.

Hinweis: Die folgenden Beschreibungen betreffen in der Mehrzahl nur die Erkennung der linearen Barcodes. Welche der Parameter bei 2D Barcode eingestellt werden, finden Sie unten im Anschluss an die Detailbeschreibungen.

"Orientierung"

Geben Sie an, in welchen Richtungen Sie Barcodes suchen wollen. Sind Sie z.B. nicht sicher, ob ein Barcode falsch herum aufgeklebt wurde, so können Sie einfach zusätzlich in entgegengesetzter Richtung suchen.

"Typen"

Bei der Schaltfläche "Typen auswählen" (s.u.) erscheint ein kleiner Dialog, in dem die gewünschten linearen Typen ausgewählt werden können. Es können mehrere Typen ausgewählt werden

"Parameter"

Hier können die voreingestellten Parameter für die Barcode-Erkennung angepasst werden (siehe Kapitel 3.2.5.4.2, Parameter).

"Schwellwert"

Barcodeerkennung wird grundsätzlich auf s/w-Bildern durchgeführt. Der Schwellwert (1-255) bestimmt den Grauwert, ab dem ein Pixel zu einem schwarzen bzw. zu einem weißen Pixel konvertiert wird.
Je höher der Wert, desto dunkler wird das Bild insgesamt. Dies gilt jedoch nur für Graustufen- und Farbbilder. Alternativ kann der Schwellwert dynamisch ermittelt werden. Über den Button "Schwellwert…" kann man den Dialog zur Konfiguration des Schwellwertes aufrufen (siehe Kapitel 3.2.5.4.3, Schwellwert).

"Erweitert"

Die "Erweiterte Suche" bietet die Möglichkeit, einige Parameter in einem engeren Rahmen während der Suche zu variieren, ohne dass ein weiterer Suchlauf gestartet werden muss. Damit erhöht sich die Wahrscheinlichkeit, dass ein Barcode erkannt wird. Die "Erweiterte Suche" kann nur bei linearen Barcodes verwendet werden (siehe Kapitel 3.2.5.4.4, Erweiterte Suche).

"Prüfsumme"

Neuere lineare Barcodetypen (z.B. Code 128) enthalten oft standardmäßig eine Prüfziffer. Diese wird im Rahmen der Barcodeerkennung überprüft und nicht zurückgeliefert. Bei Barcodetypen ohne interne Prüfziffer kann eine zusätzliche Prüfziffer berechnet und angehängt werden. Für diesen Fall haben wir einige Standardverfahren zur Verifizierung der Prüfzimmer vorgesehen, die Sie aus der Liste auswählen können.
Ist die Berechnung der Prüfsumme aktiviert, so wird diese intern durchgeführt. Passen erkannte Zeichen und Prüfziffer nicht zusammen, so wird der Barcode als falsch qualifiziert und es gibt kein Ergebnis.
Beachten Sie bitte, dass die Länge der Barcodes die gesamte Zeichenanzahl bedeutet, d.h. Nutzdaten plus Prüfziffer.

"Prüfziffer im Barcodeergebnis anzeigen"

Wird die Berechnung der Prüfsumme aktiviert, so wird diese Prüfung intern automatisch ausgeführt. Nur Barcodes mit bestandener Prüfung werden berichtet.

Im Standardfall wird die Prüfziffer nicht als Ergebnis zurückgegeben. Ist dieser Schalter gesetzt, so wird die Prüfziffer an das Ergebnis gehängt.

"Länge"

Hier geben Sie die Länge der Barcodes (Anzahl der Zeichen im Barcode) an, die Sie erkennen möchten. Barcodes mit anderer Länge werden ignoriert. Länge 0 bedeutet unbekannt.

"unbekannt -> von bis"

Wenn Sie Barcodes unterschiedlicher Längen erkennen wollen, können sie hier einen Längenbereich angeben.

"Anzahl"

Hier wird definiert, wie viele gefundene Barcodes gemeldet werden und wie mit mehreren Barcodes gleichen Inhalts auf einer Seite umgegangen werden soll.
In den allermeisten Fällen ist die Einstellung BC_MULTI die richtige Wahl, selbst wenn nur ein Barcode auf der Seite vorhanden ist. BC_MULTI liefert bei mehreren Barcodes mit identischem Inhalt nur einen dieser Barcodes zurück. BC_MULTI_MULTI müssen Sie wählen, wenn mehrere Barcodes mit gleichem Dateninhalt auf den Dokumenten sind und alle gelesen werden sollen. BC_ONE_BREAK kann in Sonderfällen gewählt werden, wenn die Erkennung sehr schnell gehen soll. Dann bricht die Erkennung sofort ab, wenn in einer Abtast-Linie auf dem Bild ein Barcode gelesen wurde. **ACHTUNG:** Durch den sofortigen Abbruch werden die gelesenen Daten nicht in anderen Abtast-Linien verifiziert und es kann zu somit Fehlerkennungen kommen, vor allem wenn die Qualität der Barcodes nicht ideal ist. Diese Option sollten Sie nur in Kombination mit einer Formatprüfung und einer Prüfziffer im Barcode in Betracht ziehen.
Die anderen Optionen (BC_ONE, BC_MULTI_ONE, BC_MULTI_BESTGUESS) sollten Sie nur nach Rücksprache mit unserem Supportteam zu wählen.

"Barcode Verdacht melden (Existenzprüfung)"

Manchmal kommt es vor, dass Belege zu hell oder dunkel gescannt sind, so dass nicht die erwartete Anzahl von Barcode-Linien sauber erkannt wird und so der Barcode nicht gelesen werden kann.
Die Anzahl der "Sollstriche" ergibt sich aus dem Barcodetyp und der Anzahl der verschlüsselten Zeichen. Wenn weniger Linien entdeckt werden, kann mit diesem Schalter in den Ergebnissen die Meldung "Barcode existiert" aktiviert werden. So wird ein "Verdacht" auf Barcodes gekennzeichnet.
Die Empfindlichkeit der Verdachtsmeldung kann in den erweiterten Parametern im Feld Prozent gesteuert werden (= Prozentsatz der Solllinien, die vorhanden sein müssen, um den Verdacht auszulösen. D.h. der Wert sollte unter 100 liegen, weil bei 100 Prozent alle Striche vorhanden müssen).

"Maximale Verdrehung"

Gibt an, bis zu welcher Verdrehung ein Barcode noch erkannt wird. Normalerweise sind 8° ausreichend. Nur wenn extremere Schieflagen zu erwarten sind, sollte ein höherer Winkel eingestellt werden. Je größer der Winkel, desto länger dauert die Erkennung.

"Intensive Suche"

Bei sehr flachen Barcodes (Seitenverhältnis ab 1:7) kann es vorkommen, dass verdrehte Barcodes durch das Suchraster rutschen. Diese Option kann auch bei qualitativ schlechten Barcodes eingesetzt werden. Mit dieser Option können Sie das Suchraster verfeinern. Achtung: Dies verlangsamt die Barcode-Erkennung!

"Konfigurationsdatei benutzen"

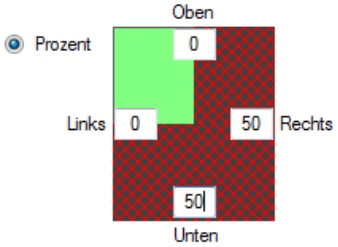
Über eine Konfigurationsdatei kann eine Bildverbesserung vor der Barcodeerkennung durchgeführt werden. Somit können Sie die Durchläufe noch feiner konfigurieren.
Bitte beachten Sie, dass die Datei nicht den Namen "qsbc.ini" trägt, da sie sonst in jedem Erkennungsdurchlauf automatisch benutzt wird.
Weitere Informationen zur Konfiguration der Bildverbesserung entnehmen Sie bitte unserem Dokument "bcTipps.pdf".

"Grafikbefehle direkt übergeben"

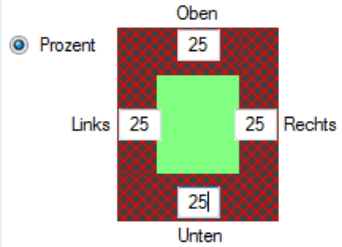
Die Befehle, die in der "qsbc.ini" stehen, können der Barcodesuche auch direkt übergeben werden. Dadurch entfällt die Notwendigkeit einer separaten Konfigurationsdatei. Die einzelnen Befehle müssen durch ein Semikolon getrennt (ohne Hochkommata) in das Eingabefeld eingegeben werden,
z.B. "FreeImage:GrayScale,8;FreeImage:Threshold,160;"

"Erkennungsbereich"

Als Erkennungsbereich kann entweder das "Gesamte Bild", ein "Bereich" in absoluten Pixelwerten oder ein prozentualer Bereich gewählt werden. Bei der Auswahlmöglichkeit "Bereich" sind die Feldkoordinaten in Pixel anzugeben. X/Y bezeichnet den linken oberen Eckpunkt des Bereichs, DX=Breite, DY=Höhe.
Bei prozentualer Angabe geben Sie die Abstände in Prozent von den jeweiligen Seitenrändern an, wobei T=Top, B = Botton, L = Links, R = Rechts. Wenn Sie prozentual einen Barcode im linken oberen Viertel suchen wollen, geben Sie Links=0, Oben=0, Unten=50, Rechts=50 ein,



für das rechte untere Viertel benutzen Sie Links=50, Oben=50, Unten=0, Rechts=0 und für einen Bereich in der Größe eines Viertels der Seite direkt in der Mitte benutzen Sie Links=25, Oben=25, Unten=25, Rechts=25.



Parameter für 2D Barcodes

Bei den zweidimensionalen Barcodetypen werden viele der Eigenschaften der Barcodes in den Spezifikationen vorgeschrieben, so dass mit der Typauswahl viele Erkennungsparameter intern optimal gesetzt werden oder für die Lesung unnötig sind.

Die Suche erfolgt bei den echten 2D Barcodes (Data Matrix, QR Code und Aztec) unter allen Winkeln (auch Zwischenwerten). Die Prüfzifferverfahren zur Erkennung von Barcodefehlern werden bei 2D Barcodes durch interne Fehlererkennungs- und Korrekturalgorithmen (ECC Error Correction Codes) ersetzt, mit denen Fehler nicht nur erkannt, sondern häufig auch automatisch korrigiert werden können.

Nach der Auswahl des Typs PDF417, Data Matrix, QR Code oder Aztec werden die nicht konfigurierbaren Einstellmöglichkeiten grau hinterlegt und können nicht verändert werden.

Die Vorgabewerte (Defaultwerte) der 2D Barcodetypen unterscheiden sich von den Defaultwerten der linearen Barcodes. Die in den meisten Fällen optimalen Defaultwerte werden bei der Typauswahl gesetzt.

Folgende Werte bleiben einstellbar:

PDF417

Orientierung (0°, 90°, 180° und 270°) und in den Parametern: Suchabstand, minimale und maximale Höhe.

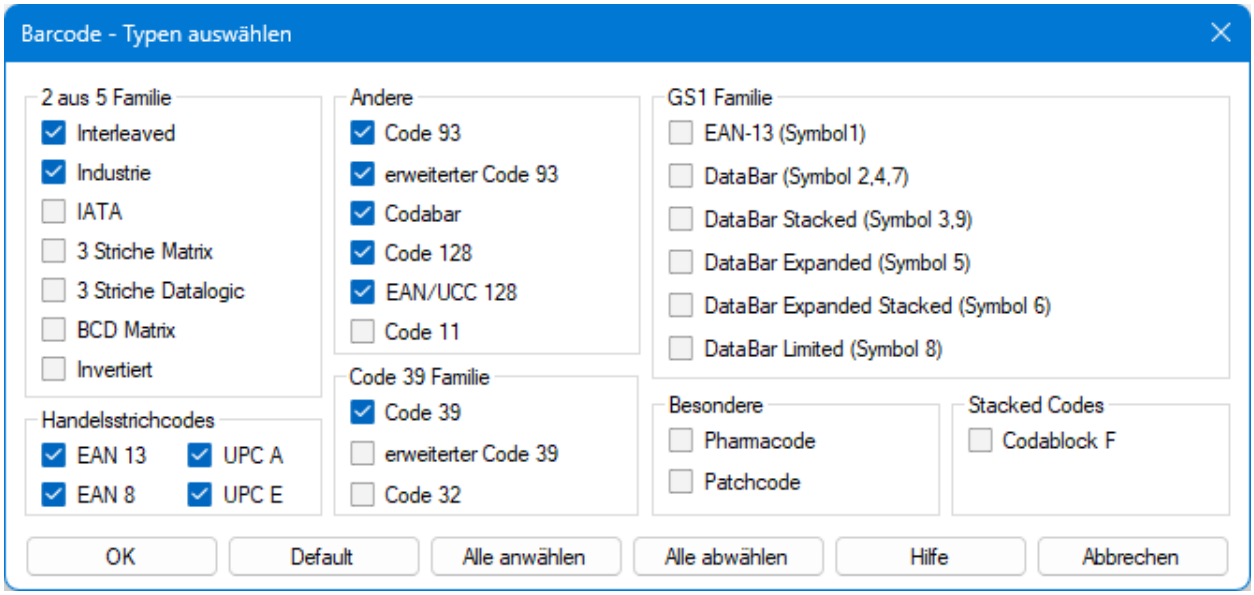
Data Matrix

In den Parametern: Ruhezone, Suchabstand, Versatz, minimale und maximale Höhe

Beim **QR Code** und dem **Aztec** Code stehen die Einstellungen überhaupt nicht zur Verfügung.

Anmerkung
Anstelle der Angabe aller obigen Parameter können Sie auch einfach alle linearen Typen in der Typauswahl anwählen und die Länge auf unbekannt zwischen 1 und 64 setzen, sowie alle Rotationsausrichtungen aktivieren. Bitte beachten Sie, dass die genaue Angabe von Typ und Länge zu besseren und schnelleren Ergebnissen führt. Falls diese bekannt sind, sollten Sie immer die entsprechenden Einstellungen wählen.

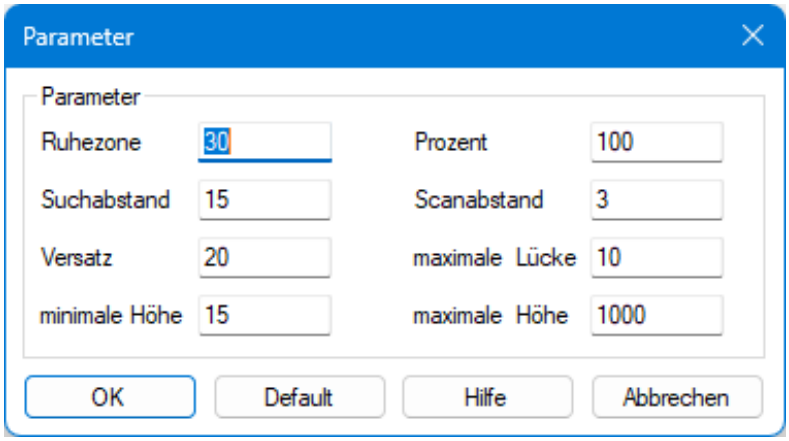
3.2.5.4.1 Barcode – Lineare Barcodetypen auswählen



Wählen Sie die linearen Barcodetypen aus, die für Ihre Dokumente in Frage kommen. Codablock F, Pharmacodes und Patchcodes können nicht gleichzeitig mit den anderen Barcodes im selben Erkennungsdurchlauf kombiniert werden.

3.2.5.4.2 Parameter

Die vorgegebenen Default-Einstellungen haben sich für lineare Barcodes in vielen Bereichen bewährt. In Sonderfällen kann aber ein Abweichen von den Werten notwendig sein, um zu Ergebnissen zu kommen.



"Ruhezone"

Anzahl weißer Pixel, die mindestens um den Barcode herum erwartet werden.
Sind die Lücken zwischen den Strichen innerhalb des Barcodes so groß, dass sie als Ruhezone erkannt werden, wird der Barcode nicht gelesen.

"Prozent"

Gibt an, wie viel Prozent eines Barcodes erkannt werden müssen, um diesen als Barcode zu melden.
Achtung, dies ist eine Sondereinstellung für Fälle, bei denen der Barcode sehr unsicher ist. Es wird "Barcodeverdacht" gemeldet, wenn Sie den Barcodeverdacht in den Einstellungen des Erkennungsdurchlaufes aktiviert haben.

"Suchabstand"

Zeilenabstand zwischen den einzelnen Abtastungen des Bildes in Y-Richtung.
Es wird automatisch auf den Scanabstand umgeschaltet, sobald die Suche genügend Striche gefunden hat. Größere Werte beschleunigen die Suche, weil weniger Linien abgetastet werden.

"Scanabstand"

Zeilenabstand zwischen den einzelnen Abtastungen innerhalb eines Barcodes in Y-Richtung.
Der Scanabstand kann bei "schlechten" Barcodes auf 1 gesetzt werden. Wenn in einer Zeile nicht mehr genügend Striche für einen Barcode gefunden werden, wird bei der Suche auf den Suchabstand zurückgeschaltet.

"Versatz"

Toleranz der Position (x-Richtung) des Barcodes beim Suchen.

"Maximale Lücke"

Erlaubter Abstand zwischen zwei erfolgreich erkannte Zeilen innerhalb eines Barcodes. Wenn dieser Abstand überschritten wird, wird der Barcode als abgeschlossen markiert.

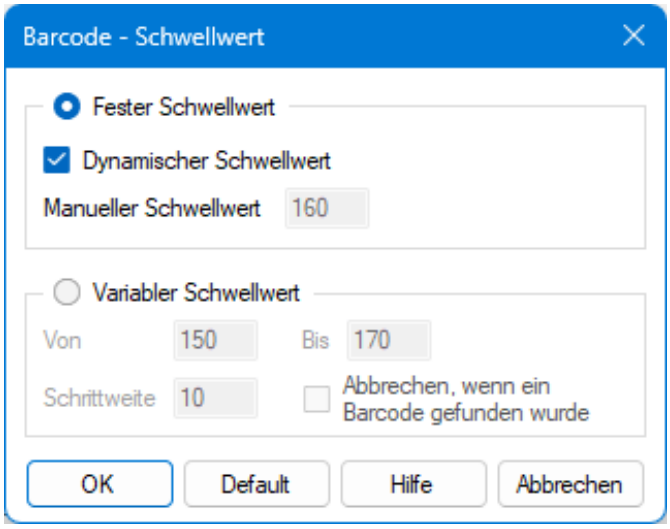
"Maximale Höhe"

Maximale Höhe die ein Barcode haben darf, um ein Ergebnis zu liefern. Größere Barcodes als die maximale Höhe werden ignoriert

"Minimale Höhe"

Die Höhe die ein Barcode mindestens aufweisen muss, um als gültiger Barcode erkannt zu werden. Bei qualitativ sehr schlechten Barcodes kann es nötig sein, diesen Wert auf einen niedrigen Wert zu setzen.

3.2.5.4.3 Schwellwert



Der Schwellwert bestimmt, ab welcher Schwelle ein Farb- bzw. Graustufenwert nach schwarz oder weiß gewandelt wird. Der gültige Wertebereich beträgt 1-255. Er kann dynamisch auf der ganzen Seite ermittelt oder manuell gesetzt werden.

Außerdem können Sie bei Dokumenten, die eine unterschiedliche Helligkeit aufweisen, z.B. weil sie mit unterschiedlichen Druckern oder Scannern erzeugt wurden, den Schwellwert innerhalb eines Erkennungsdurchlaufes stufenweise verändern, um die Erkennungsrate zu erhöhen.

Fester Schwellwert

"Dynamischer Schwellwert"

Der dynamische Schwellwert ist die Standardeinstellung. Er hat den Vorteil, dass zu hell oder zu dunkel geratene Bilder bzw. auch farbige Papiere automatisch in für die Barcodeerkennung geeignete s/w-Bilder umgewandelt werden. Bei kontrastspezifisch sehr uneinheitlichen Bildern kann es jedoch vorkommen, dass der automatische Schwellwert keine guten Ergebnisse erzielt.

"Manueller Schwellwert"

Falls der automatische Schwellwert keine guten Ergebnisse liefert, können Sie die

Umwandlung auch manuell steuern. Sie können dann selbst bestimmen, mit welchem Schwellwert ein Bild umgewandelt wird. Der Standardwert beträgt 160.

Variabler Schwellwert

"Schwellwert von"
Anfangsschwellwert, mit dem die zu verarbeitenden Bilder binarisiert werden.

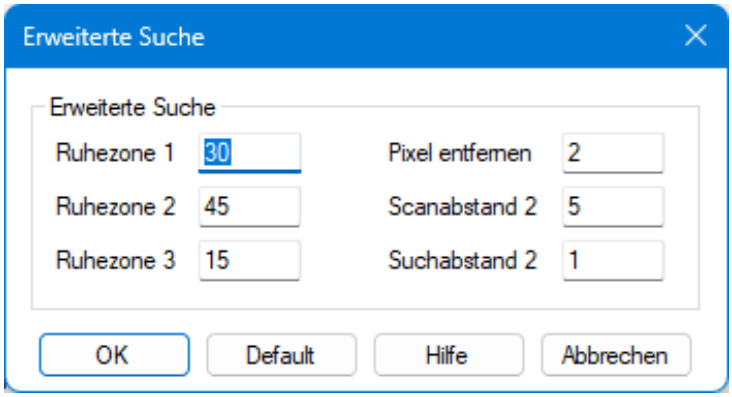
"Schwellwert bis"
Höchster Schwellwert, mit dem die zu verarbeitenden Bilder binarisiert werden.

"Schrittweite"
Der Schwellwert wird stufenweise um diesen Wert bis zum Höchstwert erhöht. Wenn ein Barcode gelesen wird, werden keine weiteren Versuche mehr durchgeführt.

3.2.5.4.4 Erweiterte Suche

Über die "Erweiterte Suche" lassen sich Suchparameter innerhalb eines Suchdurchlaufes automatisch variieren. Außerdem kann in der "Erweiterten Suche" ein zweiter Suchlauf gestartet werden, wenn beim ersten Suchlauf kein Barcode gefunden wurde.

Die "Erweiterte Suche" kann nur bei linearen Barcodes verwendet werden.



"Ruhezone 1/Ruhezone 2/Ruhezone 3"
Der Default-Wert für die Ruhezone ist 30. In der "Erweiterten Suche" können Sie bis zu drei verschiedene Ruhezonen angeben, die zur Barcodesuche herangezogen werden. Es wird grundsätzlich über alle eingetragenen Ruhezonen gesucht. Damit können also auch Barcodes mit signifikant unterschiedlichen Ruhezonen auf einer Seite gefunden werden. Bei nicht benötigten Ruhezonen setzen sie den Wert 0 ein.

"Pixel entfernen"
Häufig kommt es vor, dass insbesondere durch einen qualitativ schlechten Scanvorgang bzw. durch eine schlechte Scanvorlage kleine weiße Pixellücken in den schwarzen Balken der Barcodes entstehen. Diese kleinen Pixellücken verhindern häufig eine Erkennung des Barcodes. Die weißen Lücken (aber auch vorhandene schwarze Pixel in den weißen Lücken eines Barcodes) können bis zu einer Breite von zwei Pixeln automatisch entfernt werden. Wollen Sie keine Pixel entfernen, setzen Sie den Wert 0 ein.

"Scanabstand 2/Suchabstand 2"
Der Scanabstand gibt den Abstand zwischen den einzelnen Abtastungen des Bildes in y-Richtung an. Der Suchabstand gibt den Abstand zwischen den einzelnen Abtastun-

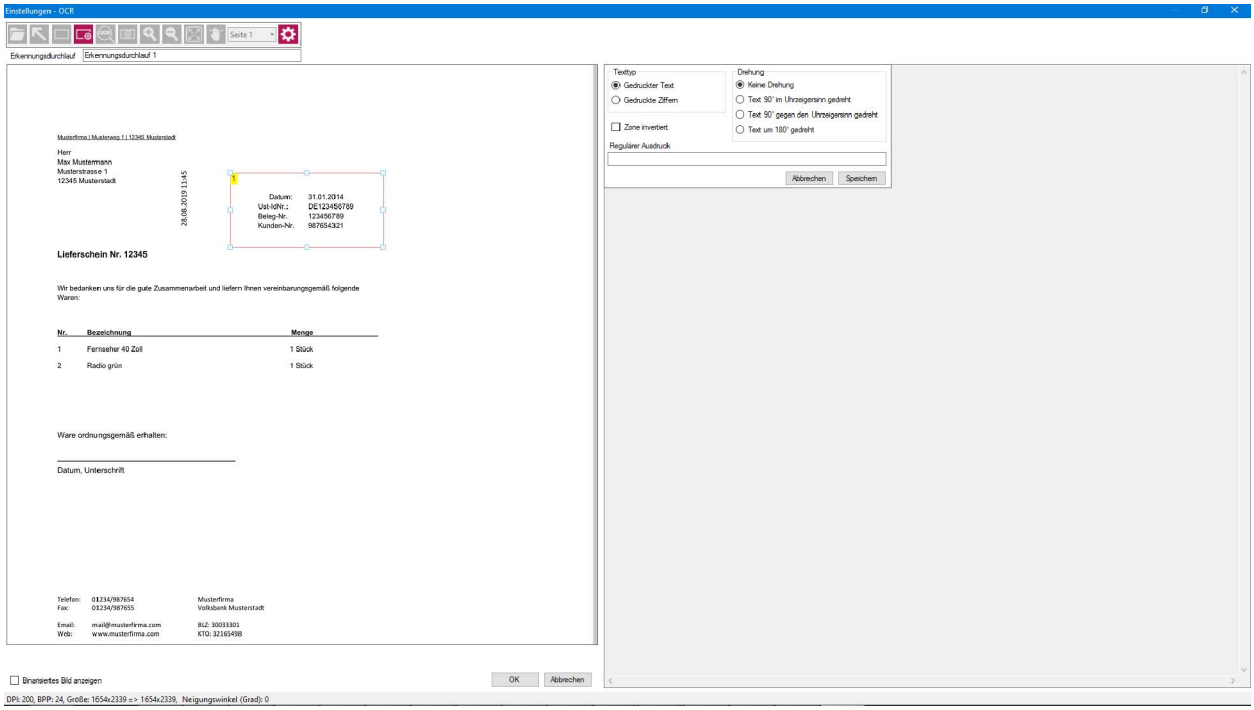
gen des Bildes in y-Richtung an, wenn ein Barcode gefunden wurde. Der erste Suchlauf wird mit dem definierten Scanabstand/Suchabstand aus dem Suchparameter-Dialog durchgeführt.
Wenn mit den angegebenen Einstellungen kein Barcode gefunden wurde, wird ein erneuter Suchdurchlauf mit dem hier angegebenen Scanabstand/Suchabstand durchgeführt. Soll der zweite Suchlauf nicht ausgeführt werden, setzen Sie beide Werte auf 0.

3.2.5.5 OCR-Einstellungen

Um zu den OCR-Einstellungen zu gelangen müssen Sie entweder einen neuen OCR-Erkennungsdurchlauf anlegen oder einen bestehenden OCR-Erkennungsdurchlauf kopieren oder bearbeiten.

In diesem Einstellungsdialog legen Sie alle Details zur OCR-Erkennung fest. Dazu zählen die Bezeichnung des Durchlaufs, die Erstellung von Zonen, in denen Text gelesen werden soll, die Leseeinstellungen der Zonen sowie die allgemeinen Parameter zur OCR-Erkennung.

Beim Öffnen des Dialoges ist dieser bis auf den Namen des Erkennungsdurchlaufes noch leer, da Sie erst ein Dokument laden müssen, auf dessen Basis Sie Lese-Zonen auf dem Dokument erstellen können, innerhalb derer dann der Text erkannt werden kann. Wenn Sie bereits Zonen für diesen Erkennungsdurchlauf erstellt haben, werden die Zonen automatisch beim Öffnen eines Dokumentes mitgeladen.



Beim Laden eines Bildes wird dieses automatisch binarisiert, d.h. in ein s/w-Bild gewandelt, da die OCR-Erkennung nur auf s/w-Bildern funktioniert. Dabei werden automatisch horizontale und vertikale Linien entfernt, die den OCR-Prozess stören würden. Zur Anzeige kommt aber zuerst immer das Original-Dokument.

Über die Checkbox "Binarisiertes Bild anzeigen" können Sie jederzeit zwischen dem Originalbild und dem binarisierten Bild hin- und herschalten und überprüfen, ob ggf. Bildartefakte die Lesung des Textes behindern.

Die Bearbeitung erfolgt über die horizontale Buttonleiste oberhalb des geladenen Dokumentes.

"**Dokument öffnen**". Öffnet ein Dokument, auf dessen Basis die Zonen zur Lesung von Text erstellt werden.

"**Auswählen**". Schaltet den Auswahlmodus für Zonen ein.

"**Zone Erstellen**". Schaltet den Modus zum Erstellen einer Zone ein.

"**Zonen-Einstellungen**". Ruft bei selektierter Zone die Zoneneinstellungen auf.

"**Text lesen**". Führt eine Lesung des Textes innerhalb der definierten Zonen durch.

"**Alle Zonen löschen**". Löscht sämtliche vorhandenen Zonen ohne Nachfrage.

"**Zoom plus**". Vergrößert die Bildanzeige.

"**Zoom minus**". Verkleinert die Bildanzeige

"**Anzeigegröße**". Stellt bei aktiviertem Zoom/Pan die originale Bildgröße/Position wieder her.

"**Pan**". Verschiebt das Bild innerhalb der Anzeige.

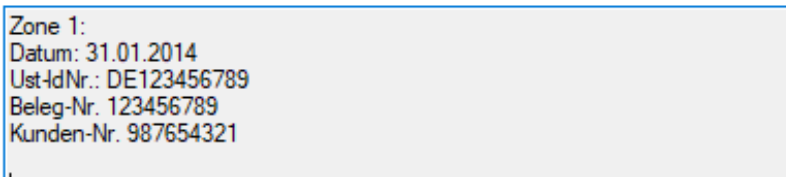
Seite 1 "**Seitenauswahl**". Auswahl einer Seite bei mehrseitigen Dokumenten.

"**Einstellungen**". Auswahl spezieller Einstellungen der OCR-Engine inkl. Sprachauswahl.

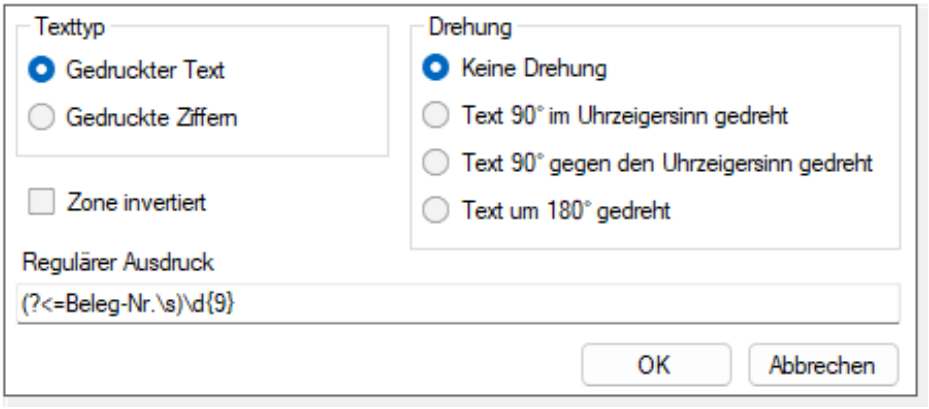
3.2.5.5.1 Die Erstellung und Konfiguration einer Zone

Um eine Lesezone zu erstellen, klicken Sie in der Toolbar auf den Button "Zone erstellen". Dann ziehen Sie mit der Maus ein Rechteck um den Bereich, aus dem während der Verarbeitung Text gelesen werden soll.

Um bereits jetzt zu prüfen, welcher Text in der Zone gelesen wird, können Sie eine OCR-Lesung über den Button "Text lesen" durchführen. Der in der Zone gelesene Text wird im Fenster rechts neben dem Bild angezeigt.



Für eine Zone können Sie verschiedene Einstellungen setzen. Um die Einstellungen aufzurufen, aktivieren Sie die gewünschte Zone mit der Maus und klicken dann in der Toolbar auf den Button "Zonen-Einstellungen.



"Texttyp"

Hier können Sie angeben, ob es sich um alphanumerischen Text oder nur um Ziffern handelt. Der Standardfall ist gedruckter, also alphanumerischer Text. Wenn Sie hier "Gedruckte Ziffern" auswählen, wird der Text ausschließlich als Zahl interpretiert.

"Zone invertiert"

Falls Sie Text in invertierten Dokumenten oder Bereichen lesen wollen. Der Bereich wird dann vor der Lesung erneut invertiert, um den Text lesbar zu machen.

"Drehung"

Wenn Sie gedrehten Text auf dem Dokument lesen wollen, können Sie die Drehrichtung angeben. Der Text wird dann vor der Lesung in die ausgewählte Richtung gedreht. Bei der Erstellung einer Zone wird bereits versucht, die Drehrichtung zu ermitteln, so dass beim Aufrufen dieser Einstellungen bereits die richtige Drehrichtung ausgewählt ist.

"Regulärer Ausdruck"

Wenn dieses Feld leer bleibt, wird der Text so übernommen, wie er in der Zone gelesen wurde. Über Reguläre Ausdrücke haben Sie eine komplexe Möglichkeit, den Text der Zone nach verschiedenen Kriterien zu Filtern.

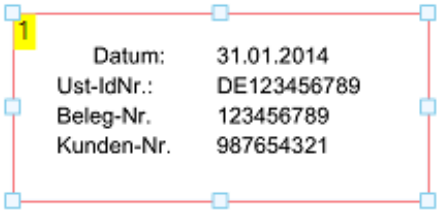
Über den Button "Speichern" können Sie die gemachten Änderungen übernehmen. Beim Abbrechen des Dialoges werden eventuelle Änderungen ohne Rückfrage verworfen.

Das Lesen von Text in einer Zone

Reguläre Ausdrücke ermöglichen die sehr feine und flexible Selektion von Text in den Zonen. Da durch den Scanvorgang die Dokumente nicht in einer identischen Position gescannt werden und die Dokumente nicht anhand von Markern ausgerichtet werden können ist es

nötig, die Lese-Zonen größer zu dimensionieren. Dies kann dazu führen, dass mehr Text als benötigt gelesen wird. Um trotzdem zum gewünschten Ergebnis zu kommen, lassen sich Reguläre Ausdrücke hier flexibel einsetzen.

Angenommen, wir haben folgende Zone auf einem Dokument erstellt, aus der wir nur die neunstellige Beleg-Nr. zur Umbenennung des Dokumentes benötigen. Eine Zone nur um die Beleg-Nr. zu ziehen verbietet sich hier aus den o.g. Gründen.



In der Zone befinden sich aber insgesamt 3 neunstellige Ziffern, so dass uns das Finden einer neunstelligen Zahl alleine nicht hilft. In diesem Fall würden wir einen Regulären Ausdruck mit Lookbehind-Funktion erstellen.

Die Lookbehind-Funktion ermöglicht das Finden von komplexerem Text, wobei bestimmte Textbereiche aus der Suchbedingung nicht mit in das Ergebnis übernommen werden. Wenn ich also die neunstellige Beleg-Nr. suche, nehme ich den Text "Beleg-Nr." mit in die Suche auf, um ihn dann aus dem Ergebnis auszuschließen.

Der Reguläre Ausdruck sähe dann so aus: (?<=Beleg-Nr.\s)\d{9}

Wir suchen nach einer neunstelligen Ziffer: "\d{9}".
Vor dieser Ziffer muss die Zeichenkette "Beleg-Nr.\s" stehen. Das "\s" steht hier für einen Tab- oder ein Leerzeichen. Der nach dem Treffer zu entfernende Text wird über die Lookbehind-Funktion entfernt: (?<=x), wobei das "x" für das auszuschließende Pattern (Beleg-Nr.\s) steht.

Das Ergebnis nach dem Lesen der Zone sähe dann so aus:

Für dieses Dokument würde das Ergebnis "123456789" übernommen und zur Umbenennung des Dokumentes führen.

Die Regulären Ausdrücke basieren auf dem Regelwerk des Microsoft .NET Framework. Weiterführende Informationen über die Sprachelemente für Reguläre Ausdrücke im Microsoft .NET Framework finden Sie unter https://docs.microsoft.com/de-de/dotnet/standard/base-types/regular-expression-language-quick-reference

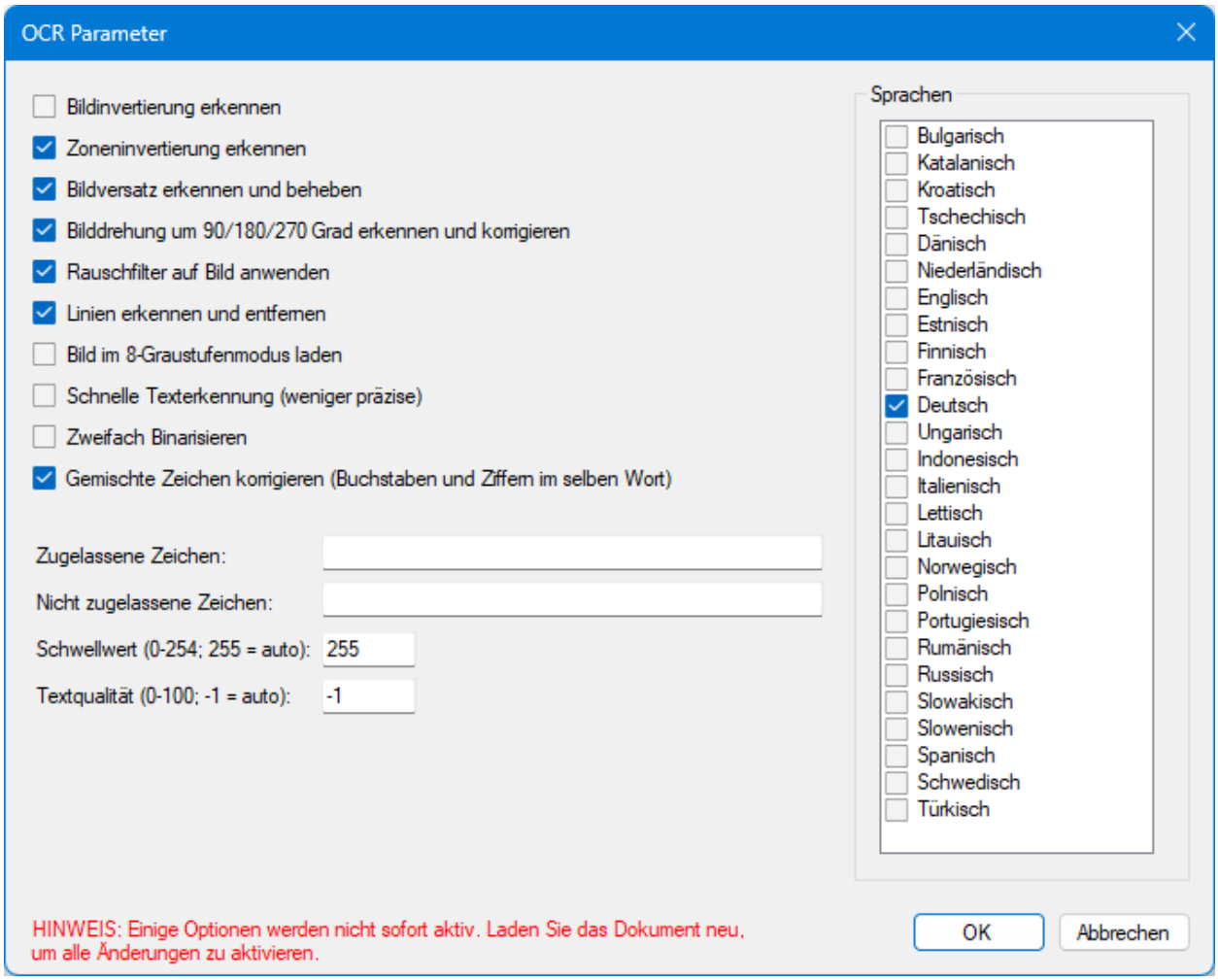
Die Behandlung von Zonen

Es können grundsätzlich beliebig viele Zonen auf einem Dokument erstellt werden, deren Ergebnisse anhand der Zonenreihenfolge dann in die Verarbeitung übernommen werden. So

wäre es z.B. auch möglich zwei Zonen zu definieren, deren Ergebnisse z.B. beim Variablen Dateinamen über die definierten Parameter weiterverarbeitet werden. Grundsätzlich werden die Ergebnisse aus der OCR-Erkennung in der Folge wie gelesene Barcodes behandelt und somit gelten alle Kriterien, die auch bei der Weiterverarbeitung von Barcode-Ergebnissen gelten.

Wenn eine einzelne Zone gelöscht werden soll, aktivieren Sie sie mit der Maus oder ziehen mit der Maus einen Rahmen um die Zone und löschen sie über die Tastatur durch das Drücken der "ENTF"-Taste. Zum Verschieben muss die Zone ebenfalls ausgewählt werden und kann dann mit der Maus oder mit den Pfeil-Tasten auf der Tastatur verschoben werden. Um dies auf mehrere Zonen anzuwenden, können Sie die gewünschten Zonen durch "STRG"+Mausklick selektieren. Alternativ können sie auch mit der Maus einen Rahmen um die gewünschten Zonen ziehen.

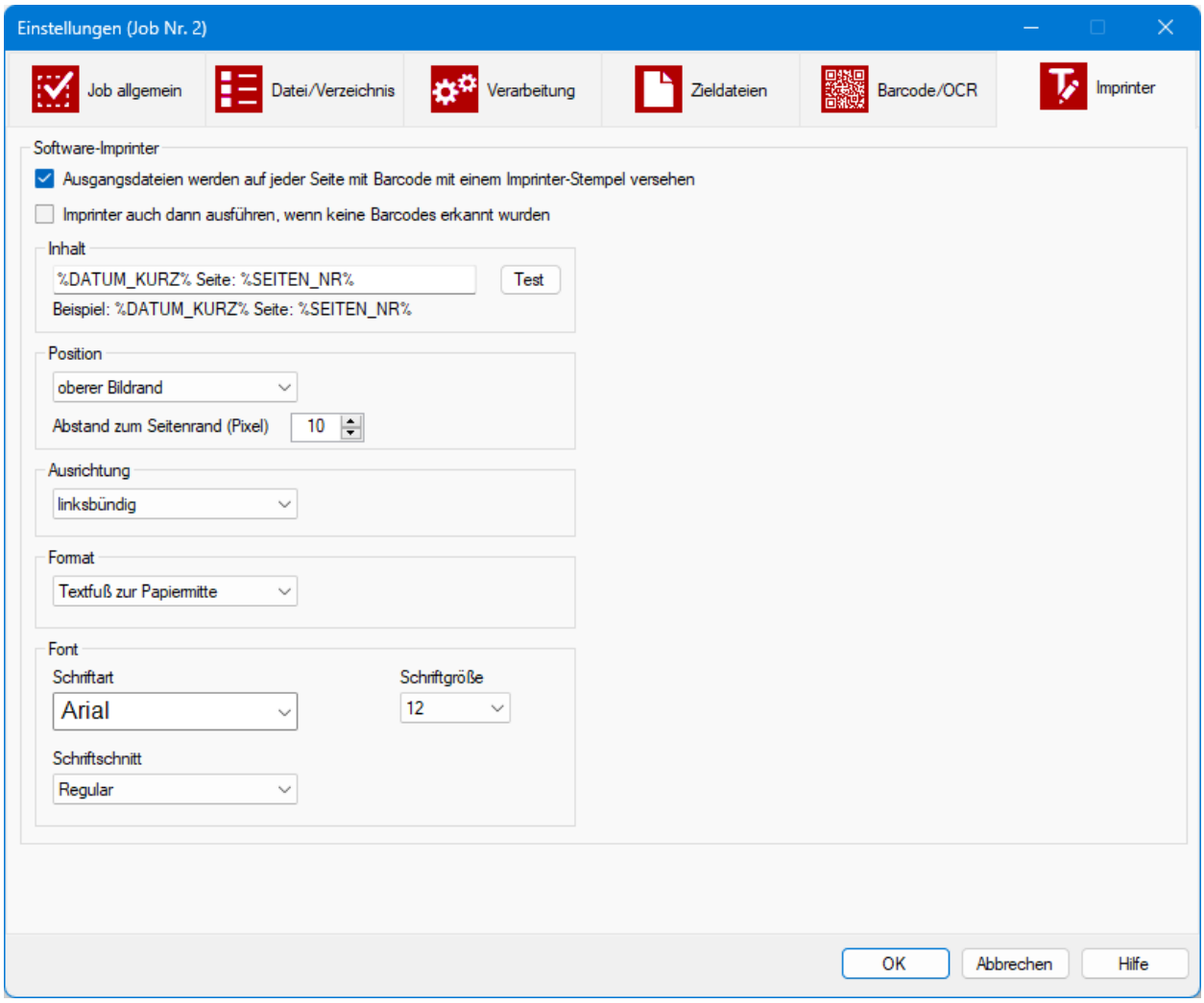
3.2.5.5.2 Einstellungen der OCR-Engine



Hier können Sie die Sprache zur Lesung des Dokumentes auswählen. Es können mehrere Sprachen parallel ausgewählt werden.

Sämtliche weiteren Standardeinstellungen der OCR-Engine sind speziell auf die Erkennung von Text in gescannten Dokumenten optimiert. Bitte ändern Sie die Einstellungen nur, wenn Sie sich über die Konsequenzen im Klaren sind. Im Zweifelsfall setzen Sie sich mit dem Support in Verbindung.

3.2.6 Registerkarte "Imprinter"



Software Imprinter

Ein "Imprinter", oft auch "Endorser" genannt, versieht gescannte Dokumente während des Scannens mit einem Aufdruck/Stempel. Dies dient dazu zu markieren, dass das Dokument bereits erfasst wurde und wird auch zur Rückverfolgung genutzt.

Ein "Software-Imprinter" oder auch "Digitaler Endorser" versieht die erzeugte digitale Kopie, also die Bilddatei, mit einer Kennung.

Die Position, das Layout inkl. der Schriftart, sowie der Inhalt des "Stempels" können definiert werden.

"Ausgangsdateien werden auf jeder Seite mit Barcode mit einem Imprinter-Stempel versehen"
: Immer, wenn ein Barcode auf einer Seite gefunden wird, wird ein Imprinter-Stempel aufgebracht.

"Imprinter auch dann ausführen, wenn keine Barcodes erkannt wurden"
: Auch auf Seiten ohne Barcode wird ein Imprinter-Stempel aufgebracht.

"Inhalt"
: Bei der Bildung des Textes, der mit dem Software-Imprinter auf verarbeitete Dokumen-

te aufgebracht wird, kann auf die folgenden Variablen zugegriffen werden, die QS-DocumentAssembler während der Laufzeit mit den entsprechenden Daten belegt:

%QUELLDATEI%	Ursprünglicher Dateiname
%BARCODE%	Der erkannte erste Barcode
%BARCODEn%	Der erkannte n-te Barcode
%BC_REGEX%	Ergebnis eines Regulären Ausdrucks auf den ersten Barcode angewandt.
%BC_REGEX_REPLACE%	Ergebnis eines Regulären Ausdrucks im ersten Barcode ersetzt.
%SEITEN_NR%	Die aktuelle Seitennummer des Quelldokumentes
%SEITEN_NRn%	Die aktuelle Seitennummer des Quelldokumentes, formatiert auf n Stellen, z.B. %SEITEN_NR3% entspricht "001"
%LFDNR%	Anzahl der Vorkommen einer Datei im Zielverzeichnis
%LFDNRn%	Die laufende Nummer formatiert auf n Stellen, z.B. % LFDNR4% entspricht "0001"
%YYYY%	Jahr vierstellig. z.B. 2010
%YY%	Jahr zweistellig, z.B. 10
%Y%	Tag im Jahr
%MMMM%	Monat Langform, z.B. "September"
%MMM%	Monat als Kürzel, z.B. "Sep"
%MM%	Monat als Ziffer mit führender Null, z.B. "07"
%M%	Monat einstellig
%DDDD%	Tag Langform, z.B. "Dienstag"
%DDD%	Tag als Kürzel, z.B. "Di"
%DD%	Tag als Ziffer mit führender Null, z.B. "01"
%D%	Tag, einstellig
%TDJ%	Tag des Jahres, z.B. 125
%HH%	Stunde, zweistellig
%H%	Stunde, einstellig
%mm%	Minute, zweistellig
%m%	Minute, einstellig
%SS%	Sekunde, zweistellig
%S%	Sekunde, einstellig
%AM/PM	Zeitangabe AM/PM, z.B. "AM"
%DATUM_KURZ%	Das aktuelle Datum im Format TT.MM.JJJJ
%DATUM_LANG%	Datum in Langformat, z.B. Dienstag, 26. November 2013
%ZEIT%	Die aktuelle Uhrzeit im Format HHmmSS

Wenn mehrere auf der Seite vorkommende Barcodes für den Imprinter herangezogen werden sollen, können Sie die Barcodes über den Parameter %BARCODEn%, z.B. %BARCODE1%, %BARCODE2%, etc. auswählen.
Die Barcodes werden in der Reihenfolge übernommen, in der Sie im Dokument gefunden werden. Diese muss nicht notwendigerweise mit der Reihenfolge der Barcodes im Dokument übereinstimmen. Sie sollten daher die Barcodes auf der Seite anhand ihrer Position sortieren lassen (siehe Kapitel 3.2.5, Registerkarte "Barcode/OCR").

Wenn mehr Barcode-Parameter als tatsächlich vorhandene Barcodes existieren, werden diese Parameter ignoriert.
Zusätzlich können Sie beliebigen Text in den Imprinter aufnehmen

"Position"
Es muss ein Seitenrand zur Positionierung ausgewählt werden.

"Ausrichtung"
Am jeweiligen Seitenrand kann der Text linksbündig, rechtsbündig oder zentriert dargestellt werden.

"Format"
Das Format gibt an, ob die Textunterkante (Textfuß) zum jeweiligen Seitenrand oder zur Bildmitte zeigt.

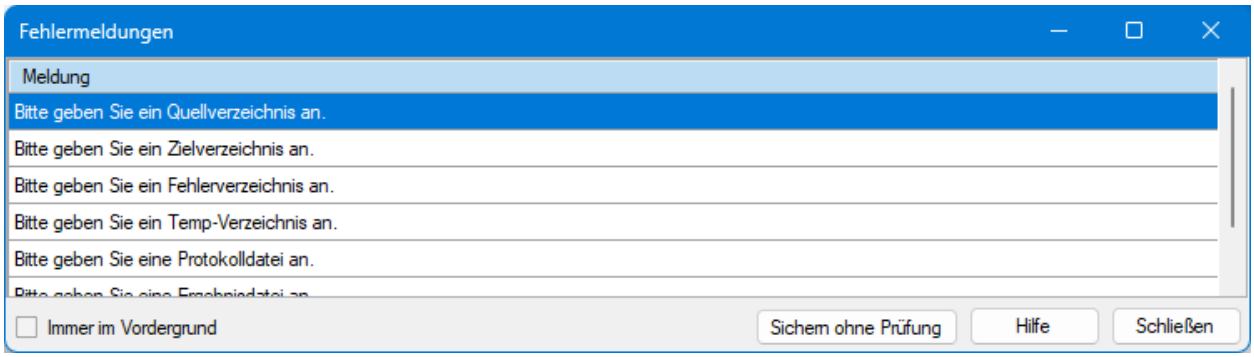
"Schriftart"
Es stehen alle im System installierten Schriftarten zur Darstellung des Textes zur Verfügung.

"Schriftgröße
Angabe der Schriftgröße, mit der der Text dargestellt werden soll. Bitte beachten Sie, dass die Zielauflösung die tatsächliche Größe des Fonts beeinflusst.

"Schriftschnitt"
Angabe des Schriftschnittes, mit dem der Text dargestellt werden soll.

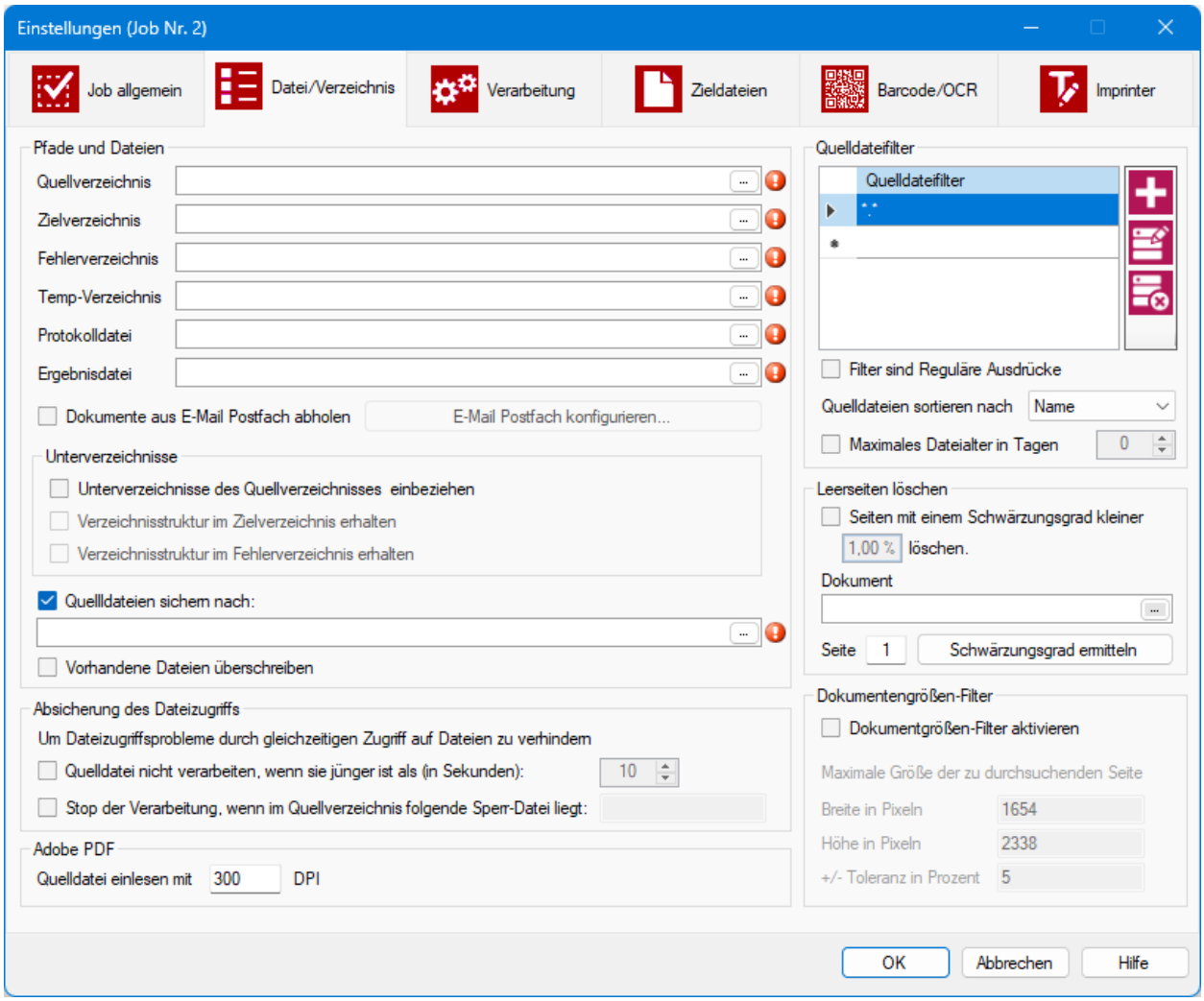
4 Plausibilitätsprüfung

Beim Speichern sämtlicher Einstellungen werden die Daten einer Plausibilitätsprüfung unterzogen. Hierbei wird im Wesentlichen geprüft, ob die Daten vollständig eingegeben wurden bzw. inhaltlich zueinander passen.



Wenn Fehler gefunden wurden, wird ein Dialog mit den Fehlermeldungen angezeigt. Aus diesem Dialog können Sie mit einem Doppelklick auf die Fehlermeldung direkt auf den Reiter springen, auf dem der Fehler aufgetreten ist. Dieses Fenster können Sie über die Checkbox "Immer im Vordergrund" dauerhaft im Vordergrund halten. Damit entfällt das lästige Suchen nach dem Fenster mit den Fehlermeldungen.

Falls Sie die Daten noch nicht komplett eingeben können oder wollen, haben Sie die Möglichkeit, die Daten aus dem Fehlerdialog auch ohne Prüfung zu sichern. Unvollständig erfasste Jobs werden in der Übersicht der Jobs als "nicht geprüft" dargestellt. Diese Jobs können nicht aktiviert und somit ausgeführt werden. Erst nach vollständiger Dateneingabe werden diese Jobs als geprüft gekennzeichnet und sind somit aktivierbar.



Die aufgetretenen Fehler werden auf dem Dialog mit entsprechenden Fehlermarken gekennzeichnet. Fährt man mit der Maus über eine Fehlermarke, wird die entsprechende Fehlermeldung in einem Tooltip angezeigt. Sie haben jetzt die Möglichkeit, alle vorhandenen Fehler zu beseitigen.

5 Einbindung des QS-DocumentAssembler in die Batchverarbeitung

Der QS-DocumentAssembler bietet Ihnen die Möglichkeit der Einbindung in Ihre Batchverarbeitung. Dies können Sie mit der Kommandozeilenversion "docasmcmd.exe" des QS-DocumentAssembler erreichen.

Der QS-DocumentAssembler unterstützt dabei einige Kommandozeilen-Parameter, bei deren Angabe die folgenden Regeln gelten:

- Groß- und Kleinschreibung ist irrelevant
- Die Reihenfolge der Parameter ist egal, bei doppeltem Auftreten desselben Parameters wird das erste Auftreten verwendet
- Jeder Schalter wird mit dem Zeichen "/" eingeleitet

/j:<Jobname>

Name des Jobs, der ausgeführt werden soll. Wenn der angegebene Name des Jobs nicht existiert, wird die Anwendung beendet.

/s:<Quellverzeichnis>

Über die Angabe eines separaten Quellverzeichnisses kann die Einstellung des Jobs übersteuert werden. Wenn das angegebene Quellverzeichnis nicht existiert, wird die Anwendung beendet.

/t:<Zielverzeichnis>

Über die Angabe eines separaten Zielverzeichnisses kann die Einstellung des Jobs übersteuert werden. Wenn das angegebene Zielverzeichnis nicht existiert, wird die Anwendung beendet.

/e:<Fehlerverzeichnis>

Über die Angabe eines separaten Fehlerverzeichnisses kann die Einstellung des Jobs übersteuert werden. Wenn das angegebene Fehlerverzeichnis nicht existiert, wird die Anwendung beendet.

Wenn die übergebenen Parameter Leerzeichen enthalten, müssen sie durch Hochkommata eingeschlossen sein.

Die Angabe des Jobnamens ist verpflichtend, alle anderen Parameter sind optional. Somit können Sie die in der Datenbank gespeicherten Verzeichnisse individuell übersteuern.

Wenn der übergebene Job durch die Kommandozeilenversion des QS-DocumentAssembler erfolgreich ausgeführt wurde, wird "0" als Rückgabewert geliefert. Kann der Job nicht ausgeführt werden, wird der Rückgabewert "1" geliefert. Weiterführende Hinweise auf die Fehlerursache finden Sie dann in der Fehlerdatei "\qsMsgLog\qsMsgLog.QS-DocumentAssembler.txt".

6 QS-DocumentAssembler Dienst

Der QS-DocumentAssembler Dienst bietet Ihnen die Möglichkeit, die Verarbeitung von Dokumenten ohne Benutzereingriff als Windows-Dienstdurchzuführen. Der Dienst bietet Ihnen dabei die folgenden Vorteile:

- Ein angemeldeter Benutzer wird nicht benötigt.
- Automatischer Neustart durch den Windows Dienst Manager
- Der Dienst wird nach einem Neustart des Servers automatisch gestartet.
- Fortlaufende Bearbeitung der Dokumente ohne Benutzereingriff.
- Der Dienst kann unter dem lokalen Systemkonto oder unter einem Benutzerkonto laufen.
- Fehlerprotokollierung in die Windows Ereignisanzeige
- Versand einer Mail im Fehlerfalle

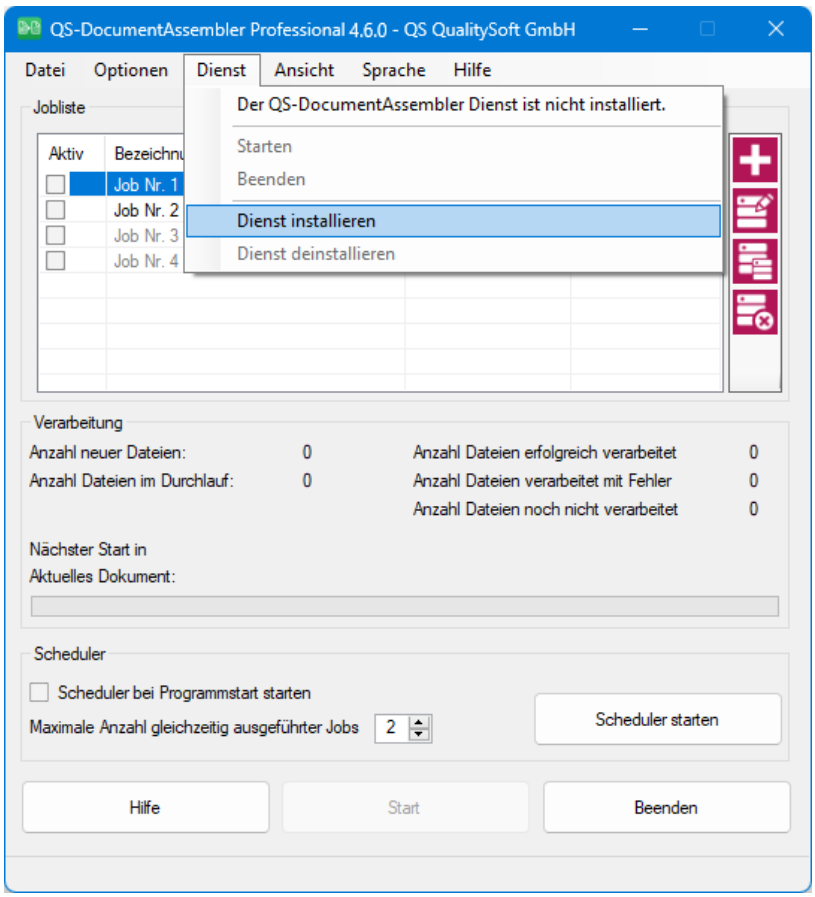
Der Dienst kann natürlich auch Dokumente von und auf Netzwerklaufwerken verarbeiten, solange die Laufwerke im Benutzerkontext des Dienstes vorhanden sind und die Pfade in UNC-Notation angegeben sind.

6.1 Dienst-Administration

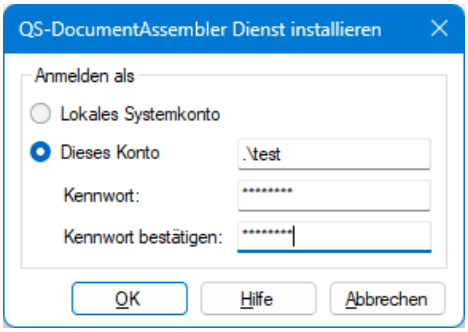
Voraussetzung für die Administration des Dienstes ist das Starten des QS-DocumentAssembler mit Administrator-Rechten.

6.1.1 Dienst-Administration über den QS-DocumentAssembler

Die Dienst-Administration kann über die Benutzeroberfläche des QS-DocumentAssembler durchgeführt werden.



Bevor Sie den Dienst nutzen können, muss er zunächst über das Menü <Dienst installieren> installiert werden.



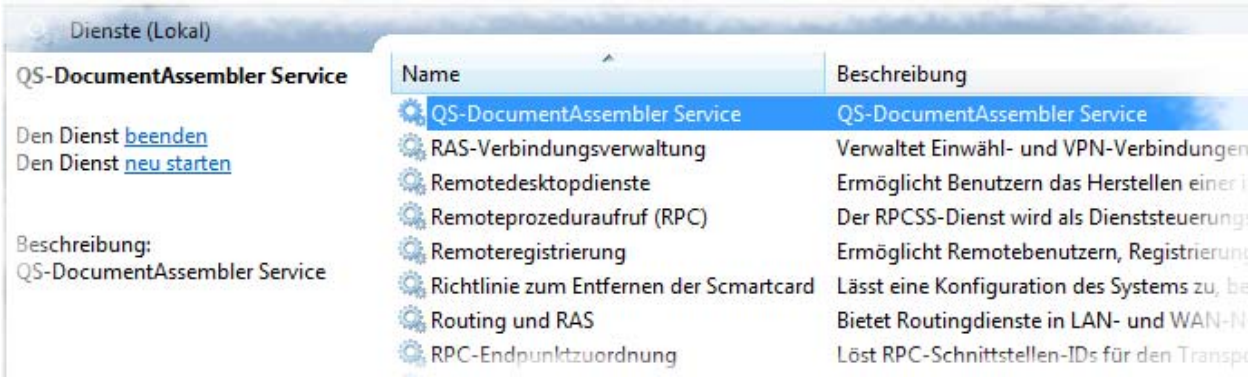
Bei der Installation müssen Sie definieren, in welchem Benutzer-Kontext der Dienst laufen soll. Er kann unter dem "Lokalen Systemkonto" oder unter einem speziellen Benutzerkonto laufen.

Wenn der QS-DocumentAssembler **Dienst** unter dem lokalen Systemkonto läuft, brauchen Sie kein Kennwort einzugeben. Wenn der Dienst unter einem speziellen Benutzerkonto läuft, muss dem Benutzernamen ein ".\" vorangestellt werden, z.B. ".\Administrator".

Wenn der Dienst unter einem speziellen Nutzerkonto läuft, muss das Konto das Recht besitzen, als Dienst ausgeführt zu werden.

Falls der Dienst auf Netzwerklaufwerke/Shares/Volumes zugreifen soll, müssen die Berechtigungen für diese Laufwerke/Shares/Volumes dem Konto zugeordnet sein.

Nach erfolgreicher Installation wird dann der Dienst der Liste der ausführbaren Dienste hinzugefügt. Hier ein Ausschnitt aus dem Windows Dienstmanager.



Über das Menü <Dienst> erhalten Sie direkten Zugriff auf den Dienst. Sie können den Dienst starten und beenden. Außerdem können den Dienst hier auch wieder deinstallieren. In der ersten Zeile des Menüs wird der aktuelle Status des Dienstes angezeigt.

6.1.2 Dienst-Administration über den Windows Dienstmanager

Alle Funktionen des Menüs <Dienst> außer der Registrierung/Deinstallation des Dienstes erreichen Sie auch im Windows Dienstmanager. Darüber hinaus können sie im Dienstmanager die Wiederherstellung des Dienstes im Fehlerfall konfigurieren.

6.1.3 Dienst-Administration über die Kommandozeile

Die Registrierung und Deinstallation des Dienstes kann auch über die Kommandozeile durchgeführt werden.

Die Registrierung erfolgt durch den Aufruf von
"C:\Windows\Microsoft.NET\Framework\v4.0.30319\installutil /i DocAsmSvc.exe"

die Deinstallation erfolgt durch den Aufruf von
"C:\Windows\Microsoft.NET\Framework\v4.0.30319\installutil /u DocAsmSvc.exe",
jeweils aus dem Programmverzeichnis des QS-DocumentAssembler.

Für beide Aufrufe benötigen Sie Administratorenrechte.

6.2 Dienst-Konfiguration

Der QS-DocumentAssembler Dienst setzt auf die vorhandenen konfigurierten Jobs auf, die für den Scheduler eingerichtet und aktiviert sind (siehe Kapitel 3.1.5, Die Erstellung eines neuen Jobs).

Der Verarbeitungsfortschritt von Jobs wird in der Oberfläche des QS-DocumentAssembler angezeigt.

Da der Dienst aber keine direkte Interaktion mit dem Benutzer durchführt, werden Optionen mit einem Interaktionsbezug wie z.B. "Anzeige einer Meldung bei Fehler" oder die Eingabe eines Präfixes bei der Verarbeitung durch den Dienst nur protokolliert bzw. nicht ausgeführt.

Wenn Sie im Dienst-Betrieb auf Netzwerklaufwerke zugreifen wollen, müssen deren Pfade in den Einstellungen, Reiter Datei/Verzeichnis zwingend in UNC-Notation angegeben werden, z.B. "\\Servername\Freigabename\Verzeichnis". Die Desktop-Version unterstützt selbstverständlich auch die "UNC-Notation". Auf lokale Laufwerke kann immer über Laufwerksbuchstaben zugegriffen werden.

Die Einstellungen eines Jobs empfehlen wir mit der Benutzeroberfläche des QS-DocumentAssembler als manuell auszuführenden Job zu testen. Nach erfolgreichem manuellem Test können Sie für diesen Job Startzeitpunkt und Intervall festlegen. Wenn der Dienst bereits läuft, wird dieser Job automatisch in die Verarbeitung übernommen.

Während der Dienst gestartet ist, kann der Scheduler nicht aktiviert werden, da sonst ein Job mehrfach gestartet werden würde und es zu Zugriffskonflikten bei den Quelldateien kommen würde. Umgekehrt kann bei laufendem Scheduler der Dienst nicht gestartet werden. Alle als manuell definierten Jobs können jedoch während des laufenden Dienstes gestartet werden, da es hier zu keinen Konflikten mit dem Dienst kommen kann.

7 Anhang

7.1 Protokolldatei

In der Protokolldatei werden Informationen über die Verarbeitung aller Belege oder nur der Belege ohne Barcode vermerkt.

Das folgende Beispiel zeigt den Aufbau der Protokolldatei. Pro Bilddatei wird eine Zeile erzeugt und die einzelnen Informationsfelder werden durch ein Semikolon getrennt.09.10.2014 09:21:02: OK; C:\DocAsm\quelle\01.tif;C:\DocAsm\ziel\021079.tif;021079;;BARCODES:1/FILECOPY:OK/FILEKILL:OK

Die Felder der LOG-Datei im Überblick:

Position	Bezeichnung	Beispiel
1	Zeitpunkt der Verarbeitung	09.10.2014 09:21:02:
2	Gesamtstatus	OK oder ERROR
3	Dateiname der Quelldatei	C:\DocAsm\quelle\01.tif
4	Dateiname, unter dem die Datei nach der Verarbeitung abgelegt wurde	C:\DocAsm\ziel\021079.tif
5	Erster erkannter Barcode auf der Bilddatei, leer falls nichts erkannt wurde	021079
6	Statusinformationen	BARCODES:1/FILECOPY:OK/FILEKILL:OK

7.2 Unterstützte Dateiformate

Der QS-DocumentAssembler unterstützt folgende Dateiformate (Auszug):

- BMP – Standard Windows Bitmap Format
- CUT – Dr. Halo/Dr.Genius Clipboard Format
- DDS – Microsoft DirectDraw Surface Format
- DIB – Standard Windows Bitmap Format
- EMF – Enhanced Windows Metaformat
- EXIF – Exchangeable Image Format
- EXR – OpenEXR Format
- FAX, G3 – Group 3 Raw Fax Format
- GIF, Interlaced GIF – Graphics Interchange Format
- HDR – High Dynamic Range Format
- IFF – Interchange Format
- J2K, J2C – JPEG-2000 Codestream
- JIF, JFIF – JPEG File Interchange Format
- JNG – JPEG Network Graphics
- JP2 – JPEG-2000 Format
- JPEG, JPG, JPE – Joint Photographic Expert Group
- JPEG progressive
- KOA – KOALA Format
- LBM – Interchange File Format-Interleaved Bitmap
- MNG – Multiple-image Network Graphics *
- PBM – Portable Bitmap File
- PBM RAW – Portable Bitmap BINARY
- PCD – Kodak Photo-CD file
- PCT, PICT, PIC – Macintosh PICT Format
- PCX – PC Paintbrush Format
- PDF – Portable Document Format
- PDF/A – Document Format for long term preservation
- PDF Multi-page – Portable Document Format
- PFM – Portable Float Map
- PGM – Portable Graymap File
- PGM RAW – Portable Graymap BINARY
- PSD – Photoshop File
- PNG – Portable Network Graphics Format
- PNM – Portable Any Map
- PPM – Portable Pixmap File
- PPM RAW – Portable Pixmap BINARY

- RAS – Sun Raster Format
- RAW camera image
- RAW memory bits – RAW bitmap
- RLE – Standard Windows Bitmap format
- SGI – Silicon Graphics Image Format
- TGA, TARGA – TARGA Image Format
- TIFF, TIF – Tagged Image Format
- TIFF Multi-page – Multi-page Tagged Image Format
- WBMP, WAP, WBM – Wireless Bitmap
- WEBP – WebP Image Format
- WMF – Standard Windows Metaformat
- XBM – X-Bitmap Format
- XPM – X Pixmap Format

7.3 Hinweise zum Adobe PDF-Dateiformat

Das Adobe PDF-Dateiformat hat sich auf dem PC mittlerweile als Standard etabliert.

Eine PDF-Datei kann unterschiedliche Komponenten wie Text, Bilder, Lesezeichen, Signaturen, etc. enthalten.

Auch im Umfeld von Dokumentenscannern wird neben den herkömmlichen Dateiformaten für Bilddateien wie BMP, TIF, JPG auch das Adobe **PDF-Format** verwendet.

In der einfachsten Form (**PDF-Typ 1**) enthalten diese PDF-Dateien **nur die Bilder** der gescannten Dokumente.

Zusätzlich kann eine automatische Schrifterkennung (**OCR**) auf die gescannten Bilder in der PDF-Datei angewendet werden, so dass die **Texte aus den Bildern** gelesen und ein "durchsuchbares PDF" entsteht (**PDF-Typ 2**). Nach diesen Texten kann dann in der PDF-Datei gesucht werden. Zu beachten ist hierbei, dass die Schrifterkennung je nach Schrift- und Bildqualität mit unterschiedlich vielen Lesefehlern behaftet sein kann. Da in der Regel keine manuelle Nachkorrektur stattfindet, können sowohl fehlerhafte Zeichen als auch Lücken im "durchsuchbaren" Text entstehen.

Außerdem gibt es natürlich noch die PDF-Dateien, die beim **"Druck" von Dokumenten** erzeugt werden (**PDF-Typ 3**). Hier sind dann die unterschiedlichen Komponenten aus den Dokumenten wie Texte, Schrifttypen, Logos, Bilder in der PDF-Datei vorhanden. Die Barcodes können hierbei als spezielle Schriftarten oder kleine Grafiken in den Seiten vorhanden sein.

Wichtig	**Eingebettete Schriftarten** Wenn Sie Schriftarten zur Erzeugung von Barcodes in Adobe PDF-Dokumenten nutzen, müssen Sie diese Schriftarten in das Adobe PDF-Dokument "einbetten", damit **QS-DocumentAssembler** die Barcode-Erkennung und Verarbeitung durchführen kann. Es reicht nicht aus, dass die Schriftart auf dem Arbeitsplatz installiert ist.

QS-DocumentAssembler unterstützt die Verarbeitung aller aufgeführten PDF-Typen.

Für die Barcode-Erkennung wird die zu lesende Seite in eine Bilddatei gewandelt, aus der der Barcode dann gelesen wird.

Hinweis: Das Lesen von Barcodes aus PDF-Dokumenten ist in der Regel möglich. Da die Programme zur Erzeugung von PDF-Dateien sehr unterschiedlich arbeiten, sollten Sie die unterschiedlichen Arten von PDF-Dateien, die bei Ihnen gelesen werden sollen, auf jeden Fall vor dem Routineeinsatz testen.

Gern unterstützen wir Sie bei Fragen zu PDF-Dateien oder zum QS-DocumentAssembler. Senden Sie uns einige Dateimuster zu, wir testen für Sie.