

LANDKREIS GÖTTINGEN



**Verbindliche Nutzung für alle Mitarbeiter*innen¹ des
Landkreises Göttingen - Fachbereich Jobcenter**

Lfd. Nr.: 2

Bearbeitung: FD 56.3 Herr Schrader, Frau Westermann

**- Bedienungsanleitung -
Drucker (Multifunktionskopierer)
Scannen mit Texterkennung (OCR)**

¹ Die in der Anleitung gemachten Angaben beziehen sich sowohl auf die männliche, weibliche als auch auf die unbestimmte Form. Zur besseren Lesbarkeit wird im Folgenden nur die männliche Form verwendet.

Inhaltsverzeichnis

1. Wozu ist eine Texterkennung (OCR) notwendig?	3
2. Ein Dokument mit Texterkennung scannen	3
2.1. SafeQ Scan auswählen	3
2.2. Scan to my email s/w und dann Scaneinstellungen auswählen	3
2.3. PDF-Menü auswählen	4
2.4. PDF-Detaileinstellung auswählen	4
2.5. Durchsuchbare PDF auswählen	4
2.6. Durchsuchbare PDF auf „Ein“ stellen	5
3. Scanfehler erkennen	5
4. Scan-Auflösung - Bildqualität - Dateigröße	6
5. Exkurs: E-Rechnung	7

1. Wozu ist eine Texterkennung (OCR) notwendig?

Scannt man ein Bild, dann erhält man nach dem Scanvorgang, je nach gewählten Einstellungen, eine Bilddatei, zum Beispiel ein JPEG. Scannt man ein Papier mit einem Text darauf, dann erhält man ohne besondere Einstellungen am Kopierer ebenfalls „nur ein Bild“ von dem Text. Das heißt, man kann den Text am Bildschirm lesen, aber man kann den Text nicht markieren und bearbeiten: Der Text kann nicht verändert werden, kein Kopieren, kein Textmarkern, keine Kommentare, keine Rechtschreibkorrektur usw.

Damit ein Text nach dem Scannen auch editiert (bearbeitet, verändert) werden kann, muss eine Texterkennungssoftware die Zeichen erst wieder in Text umwandeln. Auf Englisch heißt dieser Vorgang „optical character recognition (OCR)“, also in etwa „optische Zeichenerkennung“.

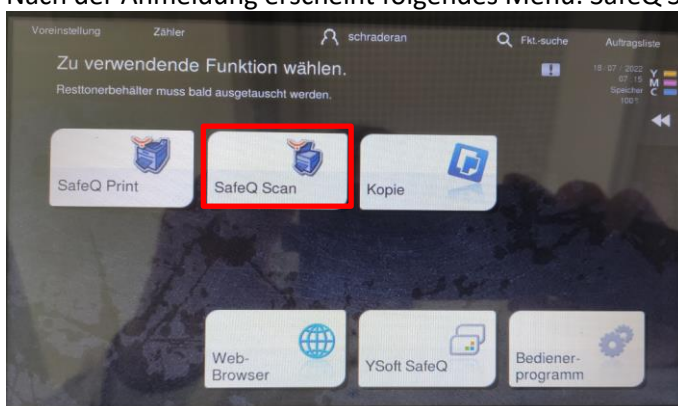
2. Ein Dokument mit Texterkennung scannen

Die Multifunktionsgeräte (Etagendrucker)² des Landkreises Göttingen können nicht nur drucken und scannen, sondern verfügen auch über eine Texterkennung, die aber im Menü gesondert ausgewählt werden muss.

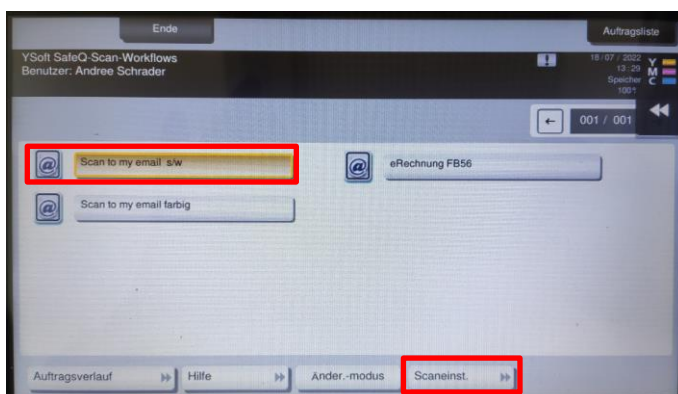
Die folgende Screenshot-Anleitung zeigt die Menübedienung am Gerät, wenn man ein Papier mit Text so scannen will, dass der gescannte Text anschließend wieder bearbeitet werden kann.

2.1. SafeQ Scan auswählen

Nach der Anmeldung erscheint folgendes Menü: SafeQ Scan auswählen



2.2. Scan to my email s/w und dann Scaneinstellungen auswählen



Es kann auch „Scan to my email farbig“ gewählt werden, aber in den meisten Fällen ist Farbe nicht notwendig.³

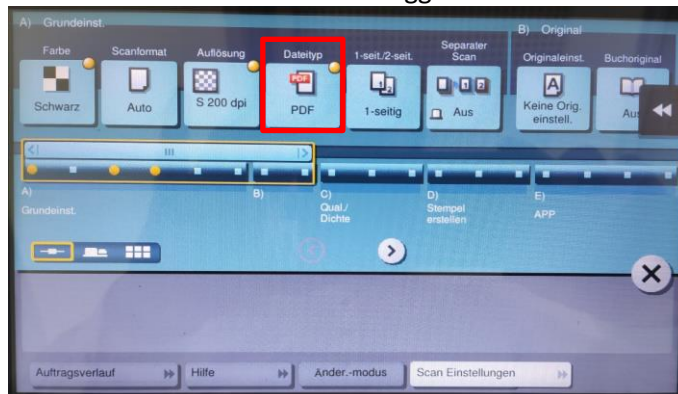
² Konica Minolta bizhub C458

³ Vgl. dazu: Kapitel 4: Scan-Auflösung - Bildqualität - Dateigröße

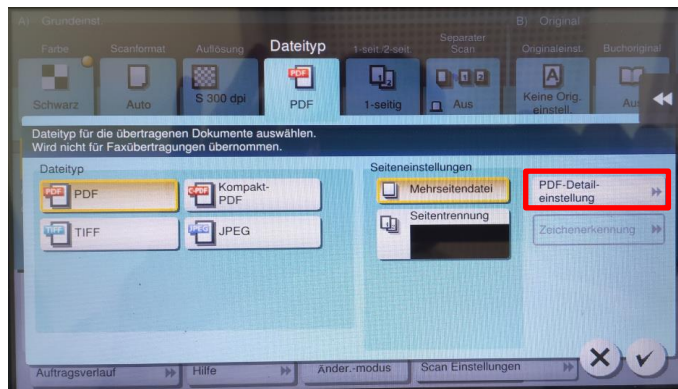
2.3. PDF-Menü auswählen

Das Feld „Dateityp PDF“ auswählen.

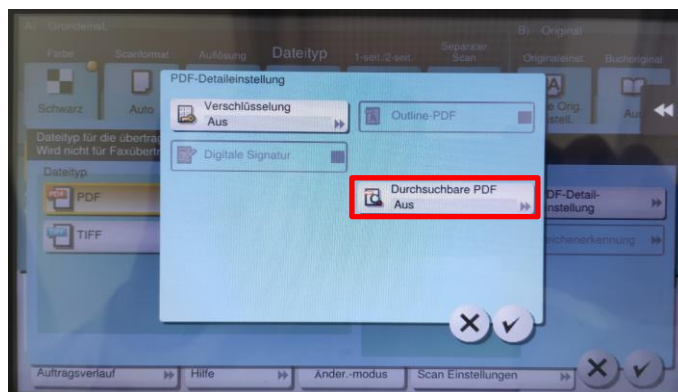
Dazu muss die obere Menüleiste ggfs. zunächst nach rechts gewischt werden.



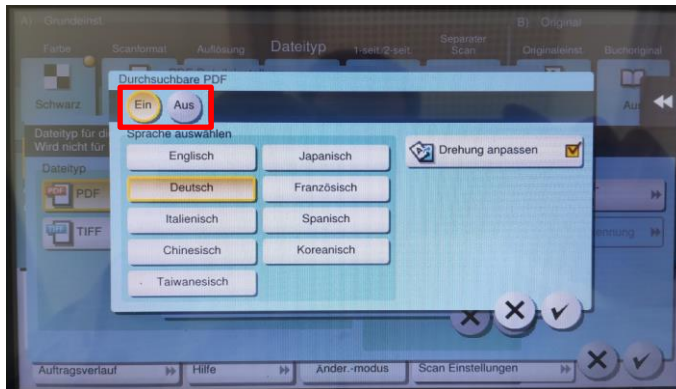
2.4. PDF-Detaileinstellung auswählen



2.5. Durchsuchbare PDF auswählen



2.6. Durchsuchbare PDF auf „Ein“ stellen



Den Knopf mit der blau beleuchteten Umrandung drücken und dadurch den Scan-to-Mail-Vorgang auslösen. Die Mail befindet sich kurze Zeit später im persönlichen Outlook-Posteingang.

3. Scanfehler erkennen

Scanfehler bei der Texterkennung (gemeint ist, dass einzelne Buchstaben oder Wörter nicht in bearbeitbaren Text umgewandelt werden konnten) erkennt man, wenn man den gescannten Text markiert und bestimmte Bereiche dabei nicht markiert werden können. Im Beispiel unten konnten die Wörter „rote Strecke“ von der OCR-Software nicht wieder in bearbeitbaren Text umgewandelt werden - sie lassen sich nicht markieren.

Harzes. Vorbei an der Sitzgruppe „Brockenblick“ erreicht man auf halbem Weg die Verpflegungsstation. Von dort geht es über den Flutberg (schöne Weitsicht ins Tal) zurück zum Ausgangspunkt.

Die rote Strecke (ca. 13 km, ±331 Höhenmeter) folgt der blauen bis zur Joseph-Quelle, biegt dann nach links ab, führt steil am Hennenberg hinauf und erreicht die Tilly-Eiche. Hier sollte man sich kurz Zeit zum Verweilen nehmen. Nach einem Steilabstieg in die Rüdershäuser Gemarkung führt uns der Weg

Wer einen gescannten Text aus einem PDF herauskopiert, um ihn in ein Word-Dokument einzufügen, muss also aufpassen, ob auch wirklich alle Wörter und Buchstaben in Text umgewandelt wurden und somit kopierbar sind. Passt man, wie im obigen Beispiel gezeigt, nicht auf, gibt es fehlende Wörter im Text, die man leicht übersehen kann. Die Wörter „rote Strecke“ fehlen im Word-Dokument (unten):

Harzes. Vorbei an der Sitzgruppe "Brockenblick" erreicht man auf halbem Weg die Verpflegungsstation. Von dort geht es über den Flutberg (schöne Weitsicht ins Tal) zurück zum Ausgangspunkt.

Die (ca. 13 km, ±331 Höhenmeter) folgt der blauen bis zur Joseph-Quelle, biegt dann nach links ab, führt steil am Hennenberg hinauf und erreicht die Tilly-Eiche. Hier sollte man sich kurz Zeit zum Verweilen nehmen. Nach einem Steilabstieg in die Rüdershäuser Gemarkung führt uns der Weg

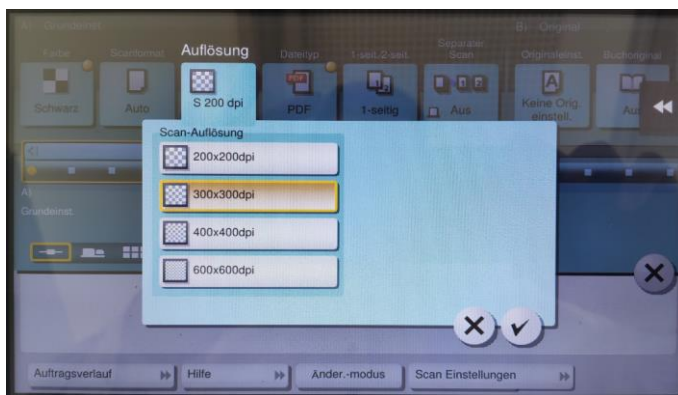
4. Scan-Auflösung - Bildqualität - Dateigröße

Die Standardeinstellung für die Scan-Auflösung ist schwarz/weiß in 200 dpi⁴. Das reicht völlig aus, damit die Texterkennung die abgebildeten Zeichen gut lesen und wieder in Buchstaben umwandeln kann. Vergrößert man ein in 200 dpi, schwarz/weiß gescanntes Bild sehr stark, dann bemerkt man, dass die Buchstaben „ausgefranst und kästchenmäßig“ aussehen.

Bedienungsanleitung

Drucker (Multifunktionskopierer)

Nur, wenn das Vergrößern eines Scans notwendig ist, z.B. für Bildausschnitte, empfiehlt es sich die Scan-Auflösung auf 300 dpi oder mehr zu erhöhen.



Um die Veränderungen der Dateigröße im Zusammenhang mit verschiedenen Scan-Einstellungen zu verdeutlichen, wurde ein Ausdruck der „OrgReg - Jobcenter-Intranet“ mit unterschiedlichen Einstellungen gescannt. Sie hat 10 Seiten, fast nur Text, eine Tabelle, die beiden Logos von Stadt und Landkreis Göttingen und nur wenig Farbe.

„Schwarz/weiß“ bedeutet, dass jeder gescannte Punkt entweder nur schwarz (vorhanden) oder weiß (nicht vorhanden) abgebildet wird.

„Graustufen“ bedeutet, dass jeder gescannte Punkt von weiß, über 256 Grautöne, bis schwarz abgebildet wird.

„Vollfarbe“ bedeutet, dass jeder gescannte Punkt farbig mit vielen Abstufungen abgebildet wird.

Scanauflösung **200 dpi, ohne** Texterkennung:

schwarz/weiß	Graustufen	Vollfarbe
0,57 MB	4 MB	5 MB

Scanauflösung **200 dpi, mit** Texterkennung:

schwarz/weiß	Graustufen	Vollfarbe
0,63 MB	4 MB	5 MB

⁴ dpi steht für engl. „dots per inch“ und sagt aus, wie viele Punkte zum Drucken oder Scannen auf einer Breite von 2,54 cm benutzt werden (Scan- oder Druckqualität). Je größer die dpi-Anzahl, desto schärfer/genauer ist die Darstellung - das bedeutet aber auch, dass der Speicherbedarf (die Dateigröße) für ein Bild/Scan größer wird.

Scanauflösung **300 dpi, ohne** Texterkennung:

schwarz/weiß	Graustufen	Vollfarbe
0,9 MB	8 MB	8 MB

Scanauflösung **300 dpi, mit** Texterkennung:

schwarz/weiß	Graustufen	Vollfarbe
0,95 MB	8 MB	8 MB

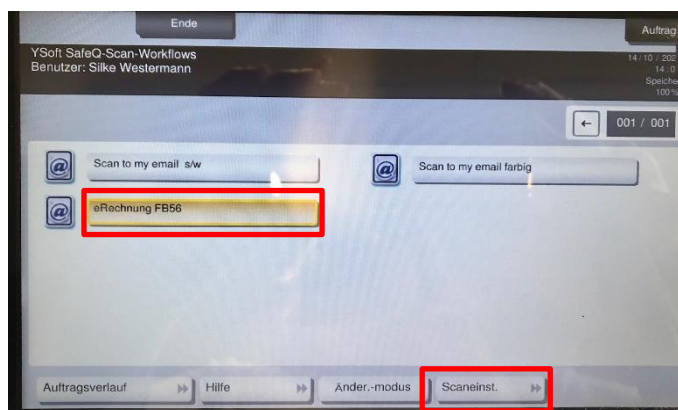
Die obigen Tabellen zeigen, dass bzgl. der Dateigröße das Scannen in schwarz/weiß der entscheidende Faktor ist. Beim Scannen mit Texterkennung steigt die Dateigröße nur unwesentlich an. Das Scannen in Graustufen oder gar in Vollfarbe führt sofort zu einer Dateivergrößerung um mindestens Faktor 6x. Die Erhöhung der Auflösung von 200 dpi auf 300 dpi führt beim Scannen zu einer weiteren Dateivergrößerung um ungefähr den Faktor 1,5x, was beim Scannen in schwarz/weiß akzeptabel ist. 300 dpi bei Graustufen oder Farbe führt gegenüber 200 dpi zu einer weiteren Dateivergrößerung um den Faktor 2x, was schon recht große Dateien entstehen lässt. Denn schließlich landen die Scans als Mail im Postfach und dort muss dann auch der benötigte Speicherplatz zur Verfügung stehen.

Um den Speicherbedarf im DMS akzeptabel zu halten und trotzdem eine lesbare Qualität zu haben, ist das Scannen mit 200 dpi in schwarz/weiß und mit Texterkennung die zu bevorzugende Einstellung.

300 dpi in schwarz/weiß verbessert die Lesbarkeit des Scans auf dem Bildschirm, die Texterkennung verbessert sich dadurch nicht spürbar. Graustufen und Vollfarbe sollten nur in begründeten Einzelfällen gewählt werden.

5. Exkurs: E-Rechnung

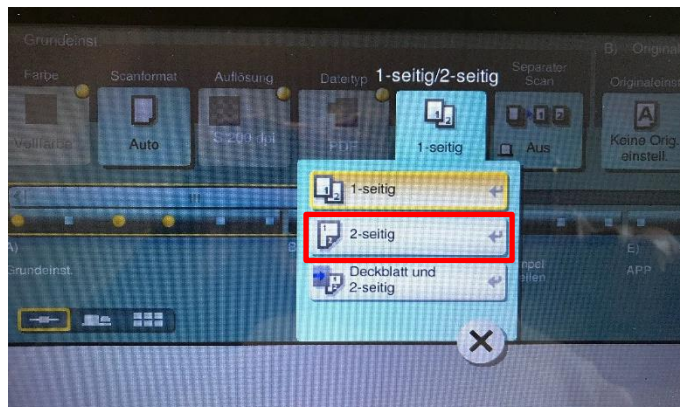
Die Multifunktionsgeräte bieten unter Scan auch die Möglichkeit „eRechnung FB 56“ an.



Diese Auswahl dient ausschließlich dazu Rechnungen, die noch auf dem Postweg eingegangen sind, zu digitalisieren und Sie in Doxis4 webCube zu importieren.

Die Grundeinstellung muss dazu nur geändert werden, wenn es sich um ein beidseitig gedrucktes Dokument handelt.

Dazu über die „Scaneinst.“ unter „A) Grundeinst.“ die Option „2-seitig“ wählen.



Sobald mit dem Knopf mit der blau beleuchteten Umrandung der Scanvorgang gestartet wird, wird das Dokument eingezogen und an Doxis4 webCube übergeben.

Achtung: wird diese Funktion versehentlich für einen „normalen“ Scan genutzt, taucht das Dokument als eRechnung in Doxis4 webcube auf und kann nur von den Hauptgeschäftsbuchhaltern wieder hieraus entfernt werden.

Freigegeben am/durch:

15.10.2024

gez. Mündemann