



# NAPS2 – Dokumente scannen mit Texterkennung

Oft möchte man Dokumente von der analogen Papierwelt in die digitale Welt übernehmen. Kein Problem – Dokument scannen, als PDF speichern, fertig. Das Resultat ist meist sehr unbefriedigend, da der Inhalt des PDFs nur aus einem Bitmap (=digitales Bild) besteht.

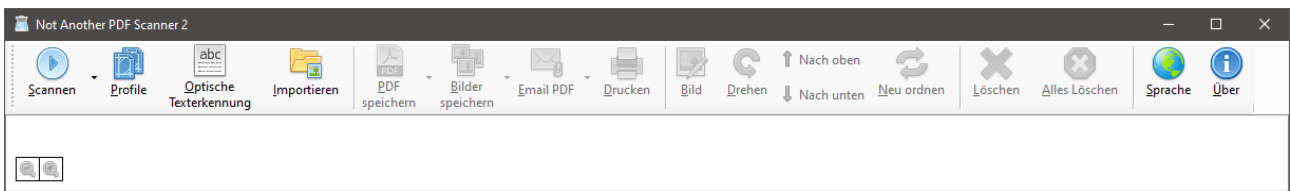
Um ein «echtes» PDF zu erhalten, muss der Inhalt mit einer Texterkennungssoftware (OCR) bearbeitet werden. Die erkannten Zeichen werden anschließend unter dem Bild mitgespeichert, sodass das resultierende PDF *durchsuchbar* wird. Natürlich kann dann auch Text über die Zwischenablage entnommen werden.

Dies alles leisten einige, oft recht teure Programme bestens, z.B. [ABBY Finereader](#) u.v.a.m.

Für schnelle Resultate gibt es von Microsoft (Konto erforderlich) die Gratis App [Office Lens](#) mit guter Texterkennung, aber leider ohne Scanner-Unterstützung (→ PDF importieren, dann erkennen).

Das kleine Open Source Programm [NAPS2](#) hat eine einfache [GUI](#), erkennt externe Scanner und arbeitet mit [Tesseract](#) (freie [OCR](#) von Google). Der Name NAPS2 – [Not Another PDF Scanner 2](#) ist ein [rekursives Akronym](#), eine unter [freier Software](#) oft verwendete Art der Namensgebung.

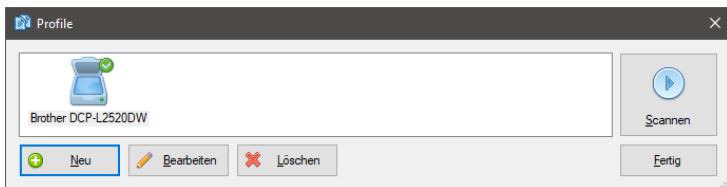
## NAPS2 starten



Die NAPS2-Oberfläche

## NAPS2 einrichten

### Scanner-Profil erstellen

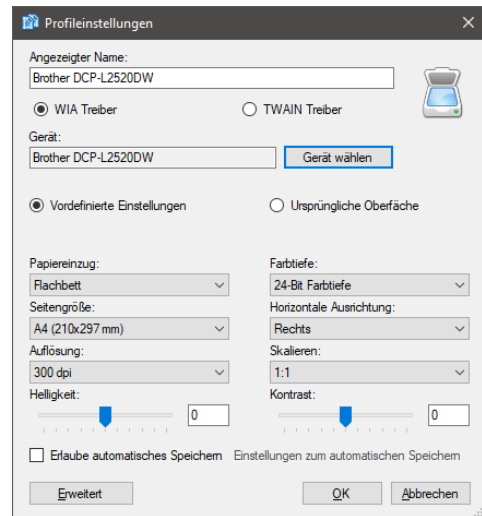


Der NAPS2 Profile-Dialog

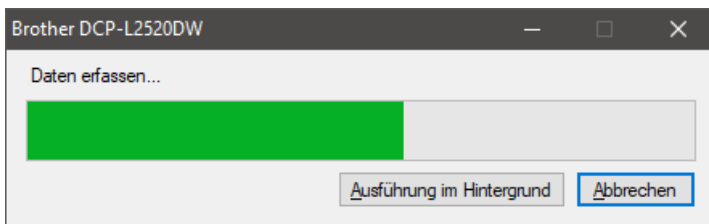
Beim Erstellen des Profils darauf achten, dass die Auflösung mindestens **300 dpi** beträgt und dass die Seitengröße auf **A4** festgelegt ist. Höhere Auflösungen liefern zwar bessere Resultate, vergrößern aber die PDFs exponentiell...

Einmal eingerichtet, erfolgt das Einlesen von Dokumenten sehr zügig.

Nun kann endlich gescannt werden...



Scanner-Profil erstellen/bearbeiten



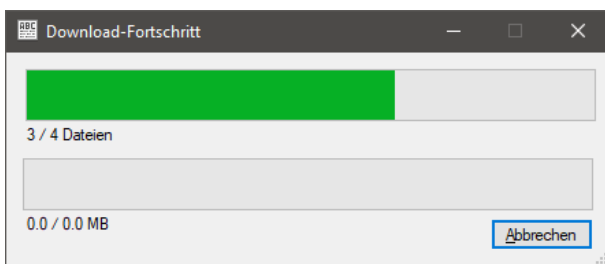
Fortschrittsbalken beim Scannen...

## OCR – Texterkennung

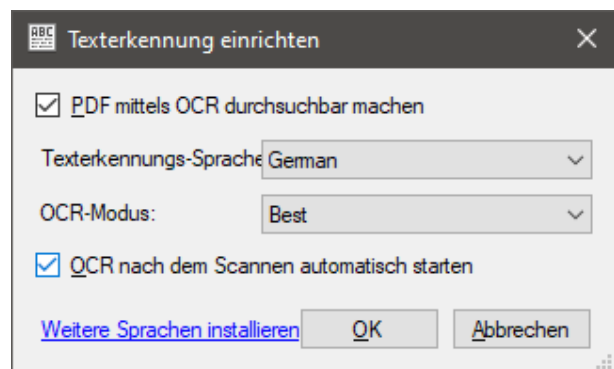
Damit die Texterkennung funktionieren kann, benötigt NAPS2/Tesseract die entsprechenden Sprachpakete (sie müssen pro Benutzer *einmal* installiert werden):



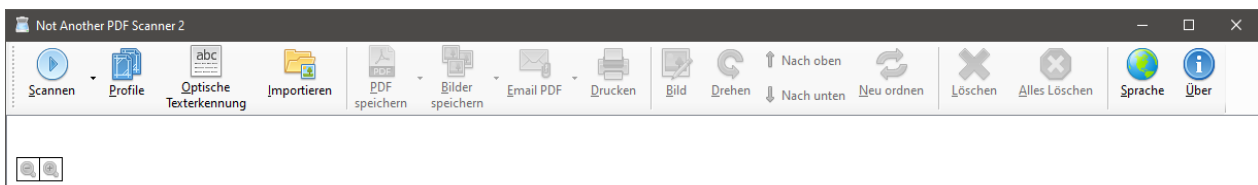
Gewünschte Sprachpakete heruntergeladen (per User)



Download & Installation – Geduld...



Dialog: Texterkennung einrichten



**TIPP:** Über «Importieren» kann auch bei bereits existierenden PDFs der Text erkannt werden (OCR).

Die Qualität der Texterkennung hängt nebst der Auflösung von vielen anderen Faktoren ab:  
Dokumentengliederung, Sprache(n), Schriftart(en) und -größen des Originals:

Übung macht den Meister!

Viel Erfolg beim Scannen mit NAPS2 wünscht

*T. Bigh*