

OCR-pdf behandling

VEJLEDNING TIL BRUGEREN

Programmet gennemgår alle pdf-filer i mappe og undermapper. Er pdf ocr-behandlet kopieres den til undermappen Pdf-ocr\<relativ sti>. Er pdf ikke OCR behandlet bliver den OCR-behandlet, og denne OCR gemmes i Pdf-ocr\<relativ sti>. Originale pdf-filer berøres ikke.

Sådan bruger du OCR-værktøjet

1. Dobbeltklik på pdf-ocr_behandling.bat (f.eks. på Skrivebordet)
2. Du bliver spurgt:

Indtast stien til mappen med PDF-filer:

3. Skriv f.eks.:

C:\Brugere\Navn\Dokumenter\MinePDFer

Sørg for at du har skrevet en **gyldig sti**

(du kan kopiere stien fra Stifinder – højreklik på mappen, vælg **Egenskaber** → kopier stien)

4. Tryk **Enter**
-

Hvad sker der?

- Alle PDF-filer i mappen gennemgås
- Hvis PDF'en allerede har tekst, kopieres den til en ny undermappe:

Pdf-ocr\

- Hvis PDF'en **ikke** har tekst, laves en OCR-version med Tesseract, som placeres i samme mappe:

Pdf-ocr\

Originalfilerne ændres **ikke**

Sådan tjekker du resultatet

Gå ind i Pdf-ocr\-mappen. Åbn filerne i din browser eller Adobe Reader og **prøv at markere tekst** med musen.

Hvis du kan markere og kopiere tekst: ☒ OCR er lykkedes.