

OCR-behandling

Elevens adgang til tekster er altafgørende for udbyttet af undervisningen. For alle, uanset læse- og skriveudfordringer, er det en fordel, at PDF-filer er elektronisk læsbare. Det vil sige, at man kan få oplæst mindre dele af en tekst eller få hele teksten læst op elektronisk. Samtidig giver det også mulighed for at redigere tekst direkte over i egne noter, lave søgning af fagord i teksten, highlighte ord og sætninger m.m.

Hvad er OCR-behandling?

OCR står for Optical Character Recognition, på dansk optisk tegngenkendelse. OCR-teknologi scanner tekst og genkender tegn ved at sammenligne tegnenes former med former, den er programmeret til at kende på forhånd. Ved OCR-behandling bliver det muligt udtrække tekst og konvertere scannede dokumenter til redigerbare, søgbare PDF-filer, Word og andre tekstdokumenter. (kilde www.ordlab.dk)

Hvad er læsbar/ikke læsbar tekst?

For at bruge oplæsningsværktøjerne i læse-skrive-teknologien er det vigtigt, at tekstfilerne er læsbare.

En læsbar tekst er i denne sammenhæng en digital tekst, hvori alle ord kan markeres. Når ordene kan markeres, er det muligt at highlighte, kopiere eller oplæse teksten. Digital tekst f.eks. Word der bliver konverteret til PDF-filer, vil ofte være læsbare, hvorimod indscannede fysiske tekster ikke vil være læsbare før de er OCR-behandlet.

OCR-behandling af ikke-læsbare filer, som ligger elektronisk tilgængelig

Der findes en række programmer som kan gøre PDF-filer og billedfiler læsbare. Der er de ordinære ordblindprogrammer som Cd-ord, IntoWords og AppWriter, men også Adobe kan bruges.

AppWriter:

I programmet AppWriter kan man OCR-behandle på to måder via knapperne OCR eller Åbn PDF.

I begge tilfælde skal der vælges den pdf-fil, der ønskes behandlet.

Vælges OCR-knappen skal der vælges, hvor filen skal gemmes. Der vil være mulighed for at vælge mellem at gemme sin OCR-behandlede fil som PDF, Word eller Text.

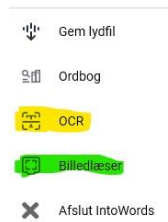
Vælges Knappen Åbn PDF, der trykkes på OCR i filen, og den gemmes ikke medmindre man aktivt gør det.



Derudover er der mulighed for at OCR-behandle en mindre mængde tekst med udklipslæser, hvis man som elev hurtigt skal have læst et afsnit af en tekst, som ikke er behandlet på forhånd. Det kan også være i multimodale tekster med både tegninger og tekst eller billedtekster, som ikke har kunnet OCR-behandles.

Intowords

PDF-filer kan behandles direkte i programmet med knappen OCR ligesom beskrevet ved AppWriter. Med billedlæser er der også mulighed for en OCR-behandling af en mindre mængde tekst / eller tekst i tegneserie /tekst i sammenhæng med et billede, som ikke har kunnet OCR-behandles.



CD-ord

I CD-ord ligger genvejene til OCR og billedlæser synlig på selve bjælken. Ligesom i de to foregående beskrivelser er der mulighed for både at OCR-behandle en tekst eller bruge billedlæser til en mindre mængde tekst.



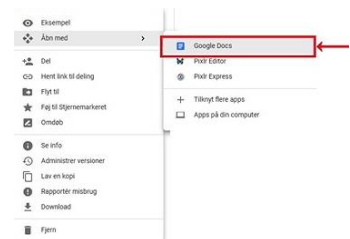
For alle tre programmer gælder, at Udklipslæser /billedlæser også vil kunne bruges på Internettet, hvor man ikke umiddelbart kan markere en tekst.

OCR-behandling via forskellige dokument /filtyper

Google Docs

PDF-filen /billedfilen gemmes på Google Drev.

Denne OCR-behandling fungerer ikke med tekst, som står i spalter og fungerer ikke udenom grafiske elementer, som ikke vil blive overført til Google Docs.



Adobe Pro

Man kan OCR-behandle i Adobe Pro. Når der åbnes en PDF-fil, vil man i højre side kunne se en funktion, der hedder Scanning og OCR.



Når der trykkes på den fremkommer der ovenover dokumentet en række muligheder. Der trykkes på Genkend tekst



OCR-behandling af fysisk tilgængelig tekst

Printer/kopimaskine:

Der kan købes licens til et OCR-program til kopimaskinen, sådan at større mængder af tekst kan scannes hurtigt ind, og OCR-behandles i scanningsprocessen. Herefter skal den indscannede tekst være til at gå til med oplæsningsprogrammerne.

OCR-behandling på mobilen (ios/android)

Der findes flere apps til IOS og Android, hvor der kan tages et billede af en tekst og få den gjort læsbar på telefonen. For de fleste apps gælder også, at de kan kopieres/deles sådan, at man på en enkel og let måde kan få dem sendt til sin computer. Derudover er der ofte flere andre funktioner, som kan være nyttige. Nedenfor bliver nævnt eksempler på nyttige apps, uden at det vil være fyldestgørende for, hvad der findes.

Google Lens

I Googles officielle app kan kameraet behandle et billede af en tekst, som kan læses op fra telefonen. Den læsbare tekst kan også kopieres til fx mail og sms. Desuden kan teksten overføres til computeren, hvis vel og mærke både computer og mobil er logget ind med samme Googlekonto.

Der vil også være mulighed for at bruge Google Oversæt, som er en del af Google Lens. Der kan der tages et billede med kameraet og få det læst op ved at trykke på tasten **vis den oprindelige tekst**.

Intowords

I appen IntoWords kan der vælges aktiviteten OCR. Der tages et billede af teksten, og herefter er den læsbare tekst markeret med en gullig farve. Efterfølgende markeres den del man ønsker oplæst. Teksten kan kopieres og deles herfra.

AppWriter:

I appen AppWriter kan der nemt og hurtigt OCR-behandles

Der er tre muligheder

- **Hurtig Skan**, hvor telefonen holdes hen over teksten, og den kan oplæses
- **Tag et billede til OCR**, hvor der tages et billede, som efterfølgende OCR-behandles.
- **Vælg billede til OCR**, hvor der ligger et billede af en tekst på telefonen, som man gerne vil have oplæst.



For iOS: mobiltelefoner med opdatering version iOS15 eller nyere

Der kan laves en livetekst med kamerafunktionen, hvor man umiddelbart efter kan få læst teksten op. Derudover kan man få teksten oversat, kopieret /delt. Endvidere er det også muligt at trykke på de enkelte ord og slå dem op i ordbog eller finde forklaringer/henvisninger på internettet.

Afrunding / anbefalinger

I forbindelse med at tekster OCR-behandles, vil

- underviseren skal altid læse teksten igennem for at sikre at teksten fremstår som ønsket og få det rettet til før det udleveres til elever. Som nævnt undervejs, vil der kunne opstå situationer, hvor der indgår tekst i forbindelse med billeder og tabeller, som man ikke kan være sikker på fremstår rigtige efter en OCR-behandling.

OCR-behandles en tekst med en blanding af billeder, faktabokse, tabeller og almindelig tekst kan specielt Word flytte rundt på formaterne.

Tekster i dansk – fx eksamenstekster, hvor der ofte er anført linjenumre ud for fx hver femte linje, kan også give udfordringer på forskellige vis.

Som vist ovenfor er der flere muligheder for at gøre tekster tilgængelige og læsbare for eleven. I undervisningssituationer, vil underviseren have det største ansvar, men elever bør også gøres bevidste om, at der er flere muligheder – fx at de altid vil have nem adgang til at få læst en kortere tekst op fra telefonen både i skolesammenhænge, men også i deres private liv.

Kreditering:

Artiklen er skrevet af Marie Jørgensen, SPS-vejleder ved UCRS, Stine Birkedal Jensen, SPS-vejleder og mentor, UCRS