



Hvordan kunstig intelligens bliver til læse- og skriveteknologier

SIGRID KLERKE, POSTDOC, INSTITUT FOR DATALOGI VED ITU, CHIEF DATA ANALYST, EYEJUSTREAD OG MARIA BARRETT, POSTDOC, DATALOGISK INSTITUT, KØBENHAVNS UNIVERSITET

Kunstig intelligens står for stort set alle nye landvindinger inden for sprogteknologi. Udviklingen går så hurtigt, at den enkelte underviser selv skal kunne vurdere relevansen af nye teknologier. Kunstig intelligens kan effektivisere nuværende læse- og skriveteknologier og synliggøre nye sider af læringsprocessen, men der er ikke noget i teknologierne, der ændrer på det faktum, at eleverne skal øve sig på at læse og skrive.

Denne artikel giver læseren en forståelsesramme for, hvor læse- og skriveteknologier *kommer fra*, og hvor sprogteknologisk forskning og kunstig intelligens tyder på, at de er *på vej hen*. Målet er at give læseundervisere et indblik i, hvordan ny sprogteknologi baseret på kunstig intelligens grundlæggende fungerer, og hvilke forventninger man realistisk kan have til den. Hermed håber vi at bidrage til at give praktikere et bedre udgangspunkt for at vurdere værdien af nye læse- og skriveteknologier. Ved at pege på, hvilke aspekter af teknologierne der er afgørende for kvaliteten, identificerer vi, hvad der er væsentligt at spørge ind til, før der investeres tid og resurser i en ny teknologi. Vores baggrund for at give dette overblik er først og fremmest baseret på erfaringer fra vores egen forskning i datadreven sprogteknologi, særligt øjenbevægelser i læsning.

På grund af den hurtige udvikling i feltet og for at værne om aktualiteten vil vi ikke her forsøge at liste specifikke teknologier på markedet, men holde os til generelle betragtninger.

Læse- og skriveteknologier skal i denne sammenhæng forstås bredt som digitale værktøjer, der bliver brugt i læse- og skrivepædagogisk praksis, og følger dermed definitionen givet i Svendsen (2017).

Læsevejlederen som gatekeeper

Læse- og skriveteknologi er ikke noget gennemudforsket område (Svendsen, 2017) og bliver det sandsynligvis heller ikke. Det skyldes først og fremmest, at den teknologiske og didaktiske udvikling ikke står stille, og at ildsjæle, opfindere og investorer kan flytte sig langt hurtigere, end forskere realistisk kan nå at gennemføre de store repræsentative normeringsstudier, longitudinelle effektundersøgelser eller pålidelige metaanalyser, som er grundstenene i videnskabelig evidens, ligesom større evalueringer af feltets samlede erfaringer naturligt kommer *efter*, ikke *før* en reform af praksis (se fx. Gandil & Olsen, 2016). Mindre studier, erfaringer, faglig sparring, interesse og nysgerrighed må derfor i praksis oftest gøre det ud for beslutningsgrundlaget for at tage en ny teknologi i brug.

Der eksisterer en fortælling om, at det værste, man kan gøre for en elev i læsevanskeligheder, er ikke at gøre noget (se fx Egmont Fonden, 2018). På den ene side kan dette 'vilde vesten' være en velkommen anledning til at eksperimenter og løbende udvikle, tilpasse og specialisere sit pædagogiske håndværk. På den anden side kan det blive en uvelkommen distraktion fra eksisterende dokumenterede undervisningsprincipper, som let overses i en blåøjet, digitaliseringsivrig jagt på en *game changer*.

Uanset hvilken side man ser på feltet fra, er det et vilkår, at når barrieren for at tilbyde noget nyt er lav, og de forskningsbaserede metoder til at verificere og validere de didaktiske resultater er både langsomme og dyre, bliver resultatet, at praktikerne selv står tilbage med ansvaret for at vurdere validiteten og legitimiteten af nye skrive- og læseteknologier på markedet.

Mainstreamteknologier og specialteknologier

Når vi tager udgangspunkt i sprogteknologisk forskning, vil det sige hele det landskab, som udforsker grænserne for, hvad computere kan gøre med sprog. Mindst tre forskellige grene af sprogteknologien har indflydelse på, hvilke læse- og skrive-teknologier der er tilgængelige i praksis: *Den første gren*, mainstreamteknologier, beskæftiger sig med helt at befri mennesker for at læse, kategorisere, oversætte og fortolke menneskesprog. Med en ingeniørteknisk tilgang bliver udviklingen hele tiden styret mod mere effektive og generelle løsninger. Eksempler på dette kan være automatisk at producere undertekster, oversættelser eller oplæsning samt søgning, sortering og opsummering i tekst-samlinger, som er for store at gennemlæse "med øjnene".

Mainstreamteknologier beskæftiger sig med helt at befri mennesker for at læse, kategorisere, oversætte og fortolke menneskesprog.

Den anden gren af sprogteknologien, som har indflydelse på markedet for skrive- og læseteknologier, ligger tættere på klassisk sprogvidenskab og er beskæftiget med, hvad computermodeller kan lære os om menneskers sprog set fra et grundforskningsperspektiv. Dette perspektiv dækker fx computermodeller, som kan være med til at afgøre, om en teoretisk model af sproget også er en realistisk kognitiv model. Fx kunne meget tidlige neurale netværk, en vigtig brik i nutidens kunstige intelligens, lære at genkende de ord, der ligner rigtige ord i et bestemt sprog (Seidenberg & McClelland, 1989). Denne model var oprindeligt udviklet til at teste den teoretiske model for ordlæsning, der kaldes konnektionistisk, men det samme underliggende princip kan i dag rette tastefejl i søgeresultater, skifte mellem dansk og engelsk stavekontrol og selv gætte, hvilket sprog maskinen skal oversætte fra.

Den tredje gren af sprogteknologisk indflydelse på læse- og skriveteknologier har et direkte didaktisk fokus og beskæftiger sig med, hvordan sprogteknologi kan hjælpe mennesker med at tilegne sig sproglige færdigheder. Specialteknologier udviklet specifikt til læse- og skriveundervisning kommer både fra den anden og den tredje gren.

De fleste af de teknologier, der allerede spiller en vigtig rolle som *kompenserende* læseteknologier, springer ud af *den første* af de ovennævnte grene af sprogteknologi; dem, der bliver udviklet først og fremmest til at give computere mulighed for at håndtere menneskesprog og løse opgaver, som ville være urealistiske at løse med mennesketimer, fx oplæsning og oversættelse af hele internettet. Kvaliteten af talegenkendelse (tale-til-tekst) og talesyntese (tekst-til-tale), OCR (optical character recognition) og endda stavekontrol og ordforslagsprogrammer bliver altså drevet mindre af markedet for specialundervisning og mere af større markeder, såsom håndfri computerbetjening og automatiseret aktiehandel. Ikke desto mindre kan disse teknologier også få stor betydning for tilgængeligheden for mennesker med særlige behov, inklusive begynderlæsere.

Det kommercielle fokus for denne udvikling er fortsat primært at erstatte klassiske sekretæropgaver, ikke læsevejlederopgaver, og det kræver, at

elever, forældre eller læsevejledere selv udvikler de understøttende læse- og skrivestrategier, som nye, alment tilgængelige teknologier muliggør (Gandil & Olsen, 2016; Svendsen, 2017; Egmont Fonden, 2018).

Den anden gren af læse- og skriveteknologier, den, der har rødder i den klassiske sprogvidenskab, inkluderer forskellige elektroniske testværktøjer som fx ordblindedetsten (Møller et al., 2014), de nationale læsetests (UVM) og optagelse af læseseres øjenbevægelser (Benfatto et al., 2016; Bingel, Barrett, & Klerke, 2018). De bygger på målemetoder udviklet i en psykologvistisk forskningstradition og er, forsimplet set, digitale erstatninger for stopur, papir, blyant, kamera og opslagstabel. Effektiviseringen består især i at gøre måling, analyse, rapportering og arkivering digital. Selvom nogle af de nyeste teknologiske tilføjelser til læsevejledernes repertoire kommer fra denne kant, bygger de på nogle ældre principper end den føromtalte gruppe af værktøjer og ligner dermed også mere papiraktiviteter, der har været brugt i praksis meget længe. Karakteristisk for denne type værktøjer er, at selvom der er oplagte fordele ved fx at kunne distribuere og analysere en test digitalt, er dens pålidelighed stadig afhængig af, at udgiveren har gennemført og dokumenteret bearbejdet med at beskrive en utvetydig testprotokol, minimere testens bias og indsamle resultater for en passende referencepopulation. Udviklingen af denne type teknologi er derfor langsommere og produkterne ofte dyrere end for de datadrevne teknologier.

Den tredje gruppe af værktøjer, den didaktiske gren af sprogteknologien, forsøger at kombinere elementer fra forskellige forskningsfelter til en form for teoretisk optimal didaktisk praksis. Elementer fra computerspil, automatisk niveautilpasning, *blended learning* med quizzer, video og kreativt skabende aktiviteter (med og uden skærmarbejde) er alle eksempler på teknologiske værktøjer, der kan have indbyggede elementer af læse- og skriveteknologi. Det er typisk praktisk umuligt at lave en forsøgsprotokol, der kan svare på, hvordan både didaktiske valg, sprogteknologiske værktøjer og andre faktorer bidrager til resultaterne i et eventuelt forsknings- eller udviklingsprojekt.

Den samvittighedsfulde praktiker på jagt efter nye evidensbaserede læse- og skriveteknologiske værktøjer står over for en markant udfordring.

Konklusionen er, at den samvittighedsfulde praktiker på jagt efter nye evidensbaserede læse- og skriveteknologiske værktøjer står over for en markant udfordring, fordi både produkterne og evidensen på forhånd må forventes at være af meget svingende kvalitet og relevans. Nogle værktøjer vil være målrettet den gruppe, som praktikeren arbejder med, andre værktøjer kan med lidt snilde bruges, selvom de er lavet til et bredt publikum. Det stærke, faglige netværk med en kultur for erfaringsudveksling og sidemandsoplæring bliver derfor let afgørende for, om en ny teknologi bliver en meningsfuld del af skolehverdagen.

Moderne kunstig intelligens lever af data

Den nyeste udvikling inden for sprogteknologi i dag sker inden for datadrevne kunstig intelligens. Kunstig intelligens er en fællesbetegnelse for systemer, der *selv* kan lære ud fra store mængder data ved hjælp af *maskinlæringsalgoritmer*, som er en samling datadrevne teknikker til at opdagge statistiske mønstre i store mængder data. I sprogteknologi er der tale om mønstre i sprogdata. Dette skal ses i modsætning til regelbaserede systemer, hvor mennesker manuelt har defineret alle beslutninger, som teknologien kan tage. Hermed er teknologien begrænset til at genkende foruddefinerede mønstre. Regelbaserede systemer kan kun vedligeholdes af eksperter inden for sprog og læring, mens systemer baseret på kunstig intelligens kan laves og vedligeholdes af ingeniører og dataloger, der ved, hvordan man hjælper *computere* med at lære noget, fx om sprog. Man kan sige, at systemet selv opdager relevante mønstre i sproget. Reelt arbejder lingvister og ingeniører oftest sammen, men pointen er, at faglig viden om didaktik nemt får laveste prioritet.

Den nyeste udvikling inden for sprogteknologi i dag sker inden for datadrevne kunstig intelligens.

Ordforslag i mobiltelefoner er et godt eksempel på, at kunstig intelligens godt nok ved, hvad der *statistisk* set er det mest sandsynlige ord, men at det ikke er det samme som, at det *for det skrivende menneske* er det korrekte eller mest meningsfulde ord. Computerens forslag er baseret på mønstre i millioner af SMS'er, men det er brugerens opgave og ansvar kun at vælge det, hun kan bruge.

Selvom grænserne for, hvad sprogteknologi kan, bliver flyttet hele tiden, vil vi stadig holde fast i, at sprogteknologi ikke ændrer på de basale spilleregler for at blive en trænet læser. På samme måde som kunstig intelligens kan gøre løbetræning mere motiverende ved at måle kroppen og fremgangen på nye måder og gøre træningen personlig, ændrer det ikke ved, at man skal løbe for at forbedre sin løbeform. Når man vælger teknologier til at understøtte læsning og skrivning, bør der stadig først og fremmest være fokus på, hvordan teknologien understøtter det, der nu engang primært skal gøres: læses og skrives.

Muligheder i data

Formålet med nogle digitale værktøjer er at gøre hidtil usynlige aspekter af skrive- eller læseprocessen synlige for eleven eller underviseren ved at registrere og analysere elevens adfærd (fx tracking af ordforslag under skrivning eller øjenbevægelser under læsning) i en detaljegrad, underviseren ikke realistisk kunne have registreret manuelt. Dette kan både være motiverende og oplysende, men værdien af sådanne teknologier bør altid vurderes ud fra deres praktiske anvendelighed. En oplagt fordel ved digitale data er, at det er relativt simpelt at lave en statistisk model baseret på denne type sprogdata uden at være sprogekspert; kunsten ligger derimod i at kunne afgøre, om modellen faktisk er nyttig, dvs. giver en øget værdi for brugerne. Dette spørgsmål kan først afgøres, når brugernes fagfelt også bliver engageret i udvikling og afprøvning.

En anden fordel ved at bruge kunstig intelligens i sprogteknologi er, at den let kan tilpasse sig sprog i forandring. Hvor et regelbaseret system fx kan være afhængig af at få en fagordsliste, kan datadrevne systemer, i princippet, bygges til løbende at lære fagord direkte fra de fagtekster, systemet trænes på. Den data, som den kunstige intelligens er trænet på, er derfor altafgørende for, hvad den lærer. Hvis man træner sit sprogteknologiske system på et *bredt* udsnit af samtidig tekst, lærer det ikke kun noget om velredigeret sprog, men også om uformelt sprog som slang, *code-switching* og sprog talt eller skrevet af forskellige subkulturer. Kunstig intelligens har med andre ord større potentiale end regelbaserede systemer, fordi det er nemmere at samle digitale sprogdata ind end at få nogen til at indkode alle regler og undtagelser på forhånd, og det er grunden til, at kunstig intelligens har vundet frem i sprogteknologien i de sidste 20 år og forventes at fortsætte med det.

Som bruger kan man ikke umiddelbart se på et system, om det er regelbaseret eller bygger på kunstig intelligens. Det kræver baggrundsviden. Men det er vigtigt at forstå forskellen, når man møder ny teknologi, da det sætter rammerne for, hvad man kan forvente af systemet, både her og nu, men også for, hvordan det kan udvikle sig fremover.

Data som begrænsning

Den vigtigste forholdsregel at tage over for systemer baseret på kunstig intelligens er, at datadrevne metoder kun virker, når der er nok af den data, som systemet skal finde mønstre i; nok læseböger, nok børnestavning, nok børnestemmer, nok ordblindestavefejl, eller hvad datagrundlaget måtte være. For de store maskinlæringsmodeller gælder tommelfingerreglen ofte: jo mere data, jo bedre model (Banko & Brill, 2001; Halevy, Norvig, & Pereira, 2009). Først når der findes adgang til at samle data ind i tilstrækkeligt store mængder, er der en realistisk chance for, at der bliver udviklet kunstig intelligens oven på det. Inden for sprogteknologi sker meget udvikling allerførst for engelsk, og det har en selvforstærkende effekt, at der laves sprogresurser, datasæt og *bench marks* til engelsk, hvorimod dansk endnu er et lavresursesprog på mange punkter. Børn er ovenikøbet, sprogtekno-

logisk set, en minoritet, fordi de skriver og læser mindre end voksne.

Børn er ovenikøbet, sprogteknologisk set, en minoritet, fordi de skriver og læser mindre end voksne.

For systemer baseret på kunstig intelligens bør man undersøge, om teknologien virker til den specifikke målgruppe. Det gælder, uanset om man vælger at bruge mainstreamteknologi eller specialteknologier lavet til sprogundervisning. Ofte vil specialteknologier lide under, at der ikke er nok tilgængelig data til at træne en god model. Hvis en teknologi er lavet til en snæver målgruppe, fx mellemtrinselever, er det heller ikke sikkert, at den virker særligt godt til ikke-typiske elever eller ældre elever med særlige udfordringer, der gør, at de læser på 3.-4. klasses niveau. På samme måde kan mainstreamteknologi virke ganske glimrende til neurotypiske, sprogstærke voksne og samtidig fungere dårligt til begynderfejl, børnestemmer, svage læsere, udtale med dialekt eller accent etc., netop fordi de statistiske modeller ikke er tilpasset målgruppens særlige sproglige adfærd. Man bør derfor forholde sig skeptisk til både specialteknologi og mainstreamteknologi baseret på kunstig intelligens, når den skal anvendes til særlige målgrupper. Den bedste og ofte eneste måde at vurdere et produkt i denne kategori er at afprøve værktøjet under så realistiske omstændigheder som muligt.

Ofte vil specialteknologier lide under, at der ikke er nok tilgængelig data til at træne en god model.

Et lovende perspektiv er muligheden for at udvikle værktøjer, der giver adgang til en kombination af deskriptive kortlægninger af elevers færdigheder samt afgrænsede, automatiske formative værktøjer til træning og test. Faren er, som beskrevet ovenfor, at den data, der bliver brugt til at bedømme elevens arbejde, ikke passer til eleven. En underviser skal derfor helst kunne se og vurdere

validiteten af den data, som en vurdering bygger på. Hvis elevens udvikling også vurderes automatisk, er det vigtigt at forholde sig til, hvad sammenligningsgrundlaget er. Der er stor forskel på at sammenligne med elevens egne data, indsamlet under sammenlignelige forhold, og at sammenligne med data fra andre elever, som har brugt systemet i en ukendt sammenhæng.

Læsevejlederen og læseren som eksperter

Når det så er sagt, er der al mulig grund til at forvente, at både stærke læsere såvel som mennesker med særlige læse- og skrivevanskeligheder vil se, at værktøjer og specialværktøjer til fx oplæsning, diktering, oversættelse og stavekontrol bliver endnu bedre. Kreative læse- og skrivestrategier vil blive ved med at opstå i mødet med andre teknologier end dem, der markedsføres som læse- og skriveteknologier.

Success i skrive- eller læseundervisningen handler ikke om, hvorvidt et værktøj er baseret på sofistikerede, statistisk maskinlæring over millioner af ord. Så længe målet er at blive en mere sikker læser, forbliver læseren selv, i samarbejde med læsevejlederen, eksperterne, der kan afgøre, om det virker, uanset hvor intelligent teknologien ellers bliver.

Referencer

Banko, M., & Brill, E. (2001). Scaling to very very large corpora for natural language disambiguation. I: *Proceedings of the 39th annual meeting on association for computational linguistics* (s. 26-33). Toulouse: Association for Computational Linguistics.

Benfatto, M. N., Seimyr, G. Ö., Ygge, J., Pansell, T., Rydberg, A., & Jacobson, C. (2016). Screening for Dyslexia Using Eye Tracking during Reading. *PLoS One*, 11(12): e0165508.

Bingel, J., Barrett, M., & Klerke, S. (2018). Predicting misreadings from gaze in children with reading difficulties. I: Tetreault, J., Burstein, J., Kochmar, E., Leacock, C., & Yannakoudakis, H. (red.), *Proceedings of the Thirteenth Workshop on*

Innovative Use of NLP for Building Educational Applications (s. 24-34). New Orleans: Association for Computational Linguistics.

Egmont Fonden (2018). *Let vejen – til uddannelse for ordblinde børn og unge*.

Gandil, T., & Olsen, M. H. (2016). *It-rygsæk til tiden*. Undervisningsministeriet: Styrelsen for Undervisning og Kvalitet.

Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*, (2), 8-12.

Møller, H. L., Arnbak, E., Petersen, D. K., Poulsen, M., Juul, H., & Elbro, C. (2014). *Teknisk rapport om ordblindetesten*. Center for Læseforskning, Københavns Universitet, og Skoleforskningsprogrammet, IUP, Aarhus Universitet.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568.

Svendsen, H. B. (2017). Et didaktiske spændingsfelt – Undervisning af unge med og i skriftsprogs-vanskeligheder, der anvender læse- og skriveteknologi. *Learning Tech 2*, 110-136. Lokaliseret d. 9. juli 2019 på: <https://tidsskrift.dk/learningtech/issue/view/7647>

UVM. *Vejledninger – om baggrund, afvikling og brug af testresultater*. Undervisningsministeriet. Lokaliseret den 7. juni 2019 på: <https://www.uvm.dk/folkeskolen/elevplaner-nationale-test--trivselsmaaling-og-sprogproever/nationale-test/vejledninger>