

CoRal Afsluttende virksomhedsrapport

RESULTATER, LÆRINGER OG VIRKSOMHEDERNES
ANVENDELSE AF CORAL

FEBRUAR 2026

01 Indledning

I en tid hvor taleteknologier udvikler sig eksplosivt globalt, har Danmark stået i skyggen, fordi vores sprogområde er for lille til, at markedskræfterne alene kan drive udviklingen. Teknologi, der forstår og gengiver talt sprog, rummer både et erhvervspotentiale og en vigtig demokratisk dimension: Den skal virke for alle borgere, også dem med dialekt, accent og/eller særlige sproglige udfordringer. Derfor bevilligede Innovationsfonden midler til CoRal (Danish Conversational and Read-aloud speech dataset), der siden 2023 har forfulgt et ambitiøst mål: At styrke udviklingen af dansk taleteknologi og bringe den op på internationalt niveau.

Projektet har siden arbejdet målrettet på at indsamle og udvikle et omfattende taledatasæt, der afspejler variationen i det danske sprog, på tværs af dialekter, accenter, alder og køn. CoRal teamet har derfor rejst rundt i Danmark, besøgt over 70 lokationer og indsamlet stemmer fra mere end 1.000 personer. Det har resulteret i op mod 1.000 timers taledata, hvoraf 330 timer er samtaler. Hver eneste stemme repræsenterer dermed en lille brik i den større vision: Taleteknologi skal kunne forstå alle danskere – ikke kun dem, der taler rigsdansk.

Der er også blevet trænet og lanceret ASR-modeller (Automatic Speech Recognition), kaldet RØST, som er frit tilgængelige for virksomheder, udviklere, forskere og myndigheder. Dataen er indsamlet i fire iterationer, således at dialektfordeling, lyd kvalitet og datatyper matcher ASR-modellernes behov og performance. Denne vekselvirkning mellem dataindsamling og modeltræning har betydet, at dataen er vokset i takt med tekniske fremskridt.

Derudover har projektet frigivet et TTS-datasæt og model (Text-to-Speech) bestående af to gange 17 timers oplæst tale fra to professionelle oplæsere. Det er et vigtigt bidrag til udviklingen af dansk talesyntese, da det kan forbedre oplæsningsløsninger i alt fra navigationssystemer til offentlige digitale tjenester.

For at sikre, at projektet ikke foregår i et teknologisk vakuum, men i tæt dialog med de aktører, der skal bruge taleteknologi, har vi i CoRal gennemført tre virksomhedsafdækninger. Her har danske virksomheder og organisationer bidraget med indsigt i deres behov, udfordringer og visioner for fremtidens taleteknologi. Samtidig har et sounding board med repræsentanter fra både private og offentlige institutioner løbende kvalificeret projektets retning. Dette har været med til at sikre, at CoRals ressourcer, datasæt og modeller afspejler reelle anvendelsesscenarier og adresserer faktiske problemstillinger i markedet.

Et åbent, demokratisk datasæt som CoRal har et enormt potentiale. Hvis virksomheder frit kan bygge løsninger, der forstår og genererer dansk tale, sænkes indgangsbarrierer markant. Det skaber et mere innovativt teknologisk økosystem, hvor både små startups, store koncerner og offentlige institutioner kan udvikle nye produkter og services. Flere open source-datasæt, -løsninger og -modeller gør det muligt at skabe alt fra mere præcise voice-bots og digitale assistenter til automatiseret dokumentation og stemmestyrede hjælpemidler. Det giver samtidig forskere og studerende en unik mulighed for at teste af og eksperimentere uden at skulle navigere i dyre licenser eller lukkede platforme.

KORT SAGT

Projektet sænker ikke blot barrierer, det har potentiale til at demokratisere hele feltet. Det er drevet af faglig styrke, nationalt engagement og en oprigtig tro på, at taleteknologi ikke kun handler om algoritmer – men om mennesker.

Indsamling af CoRal datasættet



CoRal

Hele Danmarks tale-til-tekst model

CoRal datasættet er produceret i fire iterationer, hvor der er blevet lagt vægt på 1) rekruttering af projektdeltagere på baggrund af behovsafdækning og dialekt-analyser af eksisterende data, 2) lyd kvalitet i form af optagelser foretaget i (så vidt muligt) aflukkede lokaler og med gode mikrofoner, 3) validering og annotering af den indsamlede data for at sikre bred og demografisk fordeling, og 4) udgivelse og løbende releases af både datasættet og senere RØST-modellerne.

Projektet har et særligt fokus på demokrati og diversitet i det danske sprog, ved målrettet at indsamle stemmer, dialekter og accenter på tværs af køn og alder og med deltagere i alderen fra 11 år, ledsaget af en forælder, til 97 år. Helt konkret har vi arbejdet sammen med biblioteker, kulturhuse og lignende organisationer, der har hjulpet os med både at stille lokaler til rådighed og at rekruttere deltagere via deres lokale netværk, så vores team har kunnet rejse hele landet rundt.

Det resulterede i op mod 1.000 timers taledata, som består af både oplæst tale og samtaler. Selve lydoptagelserne var delt op i to sessioner – én session med højtlesning og én session med samtale mellem to deltagere. I den første session blev hver deltager bedt om at læse foruddefinerede sætninger højt. Derudover var højtlesningen opdelt i tre runder med forskellig støj i hovedtelefoner, så vi også kunne optage forskelle i stemmeleje. Det samme gjaldt samtalesessionen, hvor to deltagere var inde og optage på samme tid. Som deltager i CoRal underskrev de en kontrakt i overensstemmelse med GDPR-lovgivning for open source, og som tak for deres tid og hjælp modtog deltagerne efterfølgende et gavekort.

For at sikre størst mulig anonymisering og adskillelse af personfølsomme oplysninger, både i datasættet og de tilhørende lydfile, er dataen blevet valideret, annoteret og transskriberet før udgivelsen. Derudover er datahåndtering, herunder overførsel og bearbejdning af taledataen, sket på sikre servere. På [Alexandra Institutet](#) kan du lære mere om vores metoder, lydsetup, transskribering m.m.

Træning og lancering af RØST-modellerne

Modeltræning og løbende releases

I CoRal har vi trænet og lanceret flere modeller, som vi samlet har navngivet RØST – efter et menneskes stemme. Modellerne er baseret på forskellige ASR-arkitekturer, specifikt Wav2Vec2 og Whisper, og vi har videreudviklet en TTS-model baseret på Chatterbox arkitekturen. Målet var og er at følge udviklingen af det tale teknologiske landskab og dermed efterspørgslen.

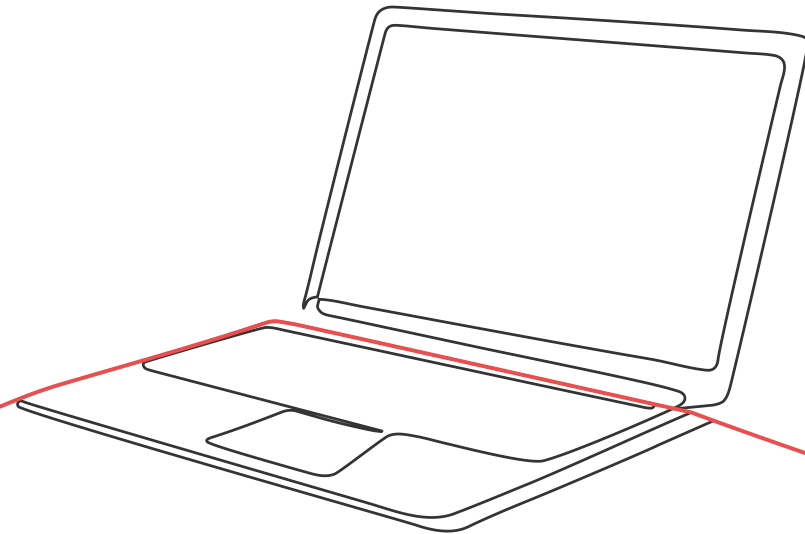
RØST-modellerne er trænet på selve CoRal datasættet, hvilket betyder, at over 468.000 individuelle ytringer, dvs. en læst eller talt sætning, er indgået i de forskellige modeltræninger. Dialekt- og accentdistributionen i taledataen spiller således en aktiv rolle i forhold til modellernes performance. Dette skyldes dels projektets fokus på demokrati og diversitet i det danske sprog, og dels de fire iterationer, hvor dataindsamlingen løbende er blevet tilpasset til de krav, særligt ASR-modellerne stillede.

Denne vekselvirkning mellem dataindsamling og modeltræning er med til at udbedre biases i feltet, da CoRals ressourcer løbende sammenlignes med eksisterende datasæt og modeller. ASR-modellernes performance måler vi mod evalueringer i testdatasæt, og deres transskriberingsnøjagtighed rapporteres som CER (Character-Error-Rate) og WER (Word-Error-Rate). Mens vi i forhold til TTS-modellen har indhentet feedback via MOS (Mean Opinion Score).

Adgang til CoRals ressourcer

RØST-modellerne kan, ligesom resten af CoRals ressourcer, frit afbenyttes og f.eks. testes af til videreudvikling af eksisterende modeller, i forbindelse med egne evalueringer, og/eller hvis du blot er nysgerrig på at vide, om det er muligt at integrere tale teknologiske services på din arbejdsplads.

CoRals datasæt og modeller er frigivet på [Huggingface](#) og via [Alexandra Institutet](#) under en OpenRAIL-D-licens, som tillader privat og kommerciel brug med nogle få undtagelser. Tag gerne direkte kontakt til os, hvis du er i tvivl om brug og licens.



Virksomhedsafdækning

For at forstå, hvordan dansk tale teknologi anvendes i praksis, samt hvilke behov og udfordringer, der præger feltet, har vi gennemført tre virksomhedsafdækninger. Den første afdækning havde fokus på at forstå problemerne ved tale teknologi på dansk og bekræftede grundlæggende barrierer såsom bias, mangel på kvalitetsdata og forskelle i dialekt- og kønsdækning. Den anden afdækning blev udført efter første udgivelse af CoRal datasættet og RØST-modellerne og havde derfor fokus på, hvorledes disse blev bragt i anvendelse hos SMV'er samt hvordan de tilgængelige ressourcer fortsat kunne tilpasses deres behov. Denne rapport præsenterer indsigter fra den tredje virksomhedsafdækning og beskriver samtidig projektet i sin helhed.

Den tredje afdækning havde tre formål: (1) at identificere hvilke aktører der anvender tale teknologi, (2) at kortlægge hvor og hvordan teknologien bruges i forskellige brancher, og (3) at belyse hvilke udfordringer – tekniske, organisatoriske og/eller sproglige – der opstår i den forbindelse. Den bidrager dermed til en bredere forståelse af, hvem der arbejder med dansk taledata og -teknologi. Aktørene spænder fra tekniske konsulent-virksomheder og softwareudviklingsmiljøer til forsknings- og uddannelsesinstitutioner samt kundeservice- og rådgivnings-funktioner. Denne bredde understreger samtidig den voksende betydning af robuste danske tale teknologiske modeller – både som teknologisk fundament og som en forudsætning for, at tale teknologi kan skalere på tværs af sektorer, brancher og siloer.

Det har dermed været et centralt fokus i foreliggende afdækning at forstå diversitet i valg af leverandører, løsninger og modeller – og hvordan organisationer navigerer i et marked i hastig udvikling. Samtidig har vi undersøgt kendskabet til CoRal i praksis, herunder i hvilken grad virksomheder allerede bruger datasættet og/eller RØST-modellerne, og hvis ikke, hvilke muligheder og fremtidige anvendelsesscenarier de ser potentiale for.

I denne afdækning har vi talt med:

- Mediacatch
- Vitec MV
- Nvidia
- Børms Vilkår
- NRGi
- N1
- Moeve
- SDU, Det Tekniske Fakultet

CoRal projektet bidrager således ikke alene med et unikt datasæt og åbne modeller, men også med indsigt i det danske tale teknologiske landskab og de behov, der skal imødekommes for at sikre, at dansk sprog forstås, understøttes og inddrages af fremtidens AI-løsninger, -modeller og -teknologier.

02 Det taleteknologiske landskab

Udviklere

Virksomheder, der udvikler taleteknologi, samt dem, der bidrager med trænings-modeller og forskning på området.

Vi har værktøjet IntoWords, som hjælper ordblinde og personer, der har skrive- og læsebesvær. Det betyder jo, at alt, der har med tekstoprettelse og tekstforståelse at gøre, er relevant for os.

VITEC MV

Det kan altså være softwarevirksomheder, teknologiselskaber eller vidensinstitutioner, der arbejder med taleteknologi.

Anvendere

Virksomheder, der anvender, træner og (videre)udvikler modeller, services og løsninger. Det kan være til intern brug og/eller som del af deres produktportefølje.

Vi har nærmest fra starten lavet, trænet og udviklet på tale-til-tekst-modeller. Vi bruger det som et middel til et andet mål, som er at servicere vores kunder.

MEDIACATCH

Ud fra deres arbejde med taleteknologi kan virksomhederne også tilbyde rådgivning om anvendelse og implementering.

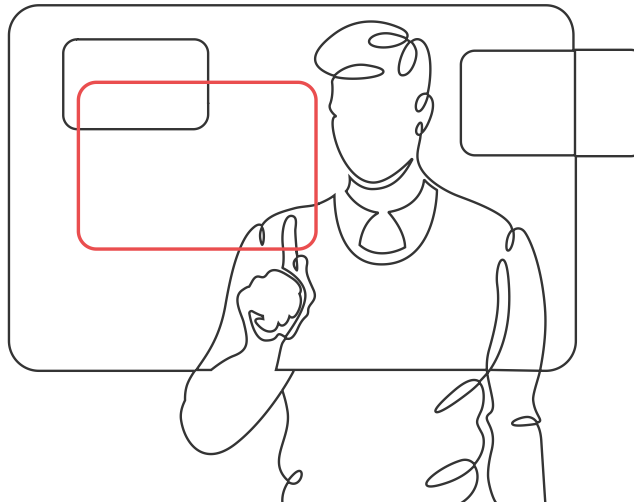
Aftagere

Virksomheder eller (offentlige) organisationer, der aftager taleteknologiske services.

Jeg fandt [CoRal] webinarer interessant, fordi jeg kort inden fik til opgave at undersøge, hvad omfanget ville være for at udvikle en transskriptions-service til interne møder. Vi har nemlig oplevet utilfredshed med transskriberingskvaliteten af Teams.

NRGi

Ofte opstår interessen for taleteknologi ud fra interne udfordringer, behov, ønsker og/eller en digital forestillingsevne om procesoptimering.



Udvikling af taleteknologi

Taleteknologi er fuldstændig, vanvittig vigtigt nu til dags.
MEDIACATCH

Vi vil egentlig gerne køre lokalt på brugernes egne devices, og der er muligheder hist og her, også med Microsoft, som vi bruger. Men det er meget vigtigt for os, at der ikke er lange arme fra et amerikansk selskab ned i vores løsninger, og det kan være en udfordring.
VITEC MV

Det er ikke så skræmmende og fremtidsagtigt længere.
SDU

På tværs af brancher og siloer oplever vi en øget bevidsthed omkring taleteknologi. Dette medfører ikke blot en voksende interesse og stigende efterspørgsel, men også en større digital forestillingsevne.

Flere virksomheder peger på, at udviklingen sker i takt med, at flere og flere ser potentiale frem for den automatiske skepsis, der måske tidligere prægede feltet. Markedet afsøges nu bredt og målrettet, der gennemføres stadig flere pilotprojekter, og teknologierne testes løbende på modenhed, kvalitet og forretningsværdi.

Det er især branchespecifikke behov, arbejdsgange og fagsprog, der bestemmer de krav, der stilles til taleteknologi. Nogle virksomheder ønsker bedre transskription af interne møder, andre har fokus på optimering af kundeservice, og igen andre er interesseret i mere ensartet dokumentation og sagsbehandling. Kort sagt er det, der tidligere var utænkeligt i en dansk kontekst, blevet til realitet.

Mulighederne er mange, men udviklingen af taleteknologi resulterer samtidig i en øget kritisk bevidsthed om anvendelsen, en voksende opmærksomhed på datakvalitet samt et ønske om digital suverænitet. Nogle virksomheder understreger eksempelvis, at de ikke ønsker at være bundet til én bestemt leverandør, model og/eller teknologi.

Voksende interesse og stor forventet værdi

Virksomhederne ser taleteknologi ikke kun som en teknologisk løsning nu og her, men som et område i hastig udvikling, hvor potentialet rækker langt ud over de enkelte use cases.

Det er ikke specifikt lige for denne her branche, men jeg tænker sådan mere generelt, at alt omkring taleteknologi bliver stort. Fordi kan du oversætte tale til tekst og smide det ind i en computer, kan du bruge det som prompt på alle mulige måder.

N1

Det handler for mange virksomheder især om digitalisering og effektivisering af processer, hvor taleteknologi bliver en del af arbejdsgangen, hvilket i sidste ende frigiver tid til andre opgaver. Derudover er den rimelig nem at anvende og forholdsvis billig.

Inden for brancher, hvor der er krav om meget dokumentation, og hvor der er mange håndholdte devices, kan det være en fordel med tale-til-tekst. Fordi kvaliteten er så høj nu, at folk slipper for at stå og taste med én finger ad gangen, så det kan faktisk øge produktiviteten at have ordentlig tale-til-tekst-teknologi.

VITEC MV

Mulige anvendelsesscenarier spænder derfor bredt, både internt i organisationen og på tværs af brancher. Målet er typisk, at det skal lette medarbejdernes hverdag, skabe mere tid til faglig fordybelse og styrke kontakten til kunder, borgere eller kollegaer.

ANVENDELSESSCENARIER FOR TALETEKNOLOGI

- Transskription af møder, medieindhold etc.
- Analyse og segmentering af brands
- Diktering, journal- og referatskrivning
- Udkast til handleplaner
- Live undertekstning
- Talesyntese
- Oversættelse og tolkebistand
- Understøttelse af sagsbehandling
- Effektivisering af informationssøgning



Erfaringer med taleteknologi

Anvendelse og udvikling af taleteknologi

Virksomhederne har efterhånden gjort sig en række erfaringer med anvendelsen af taleteknologi og dens hidtidige udvikling.

Tidligere var det meget kvaliteten af det, som var en roadblock, nu er det lidt sådan, når først brugerne forstår, at de skal indtale en tekst og blive ved med at tale, altså hvordan teknologien skal bruges, så giver det faktisk ret høj kvalitet.

VITEC MV

Kvaliteten er blevet bedre, modellerne er blevet mere diverse, og stadig flere brugere er i højere grad fortrolige med dem. Samtidig peger flere virksomheder på udfordringer, særligt i forhold til retskrivning, hastighed og kontekstforståelse.

Det er cirka hver tiende ord, der staves forkert, og det er noget skidt, hvis de ord, der staves forkert, er de vigtige, for eksempel brands. Der er modellen, som vi bruger, bare ikke altid helt intelligent nok, så kontekstviden er en udfordring for os stadigvæk. Whisper og den type modeller er lidt bedre til det, men de er så kontra vores langsommere, og uegnet til live undertekster, i hvert fald i udgangspunktet. Så der vil altid være klassiske ingeniørmæssige trade-offs.

MEDIACATCH

Erfaringer fra praksis viser, at der fortsat er udfordringer med taleteknologi, og det vil der måske altid være. Ikke desto mindre er der overordnet set stor tilfredshed blandt virksomhederne, og mange ser med spænding frem til udviklingen på dette område.

Valg af leverandører

Nogle virksomheder er stadig i gang med at afsøge markedet, muligheder og potentialet for taleteknologi. Her er faktorer som økonomi, tidsomfang, GDPR, datasikkerhed, kvalitet og type af data (herunder retskrivning, lyd kvalitet, fagsprog etc.) afgørende, når der skal vælges en passende leverandør.

Vi startede med tanken om at lave det selv, men så screenede vi markedet og fandt to-tre firmaer, der var i gang med det, og som havde noget kørende, der virkede fint. Så valgte vi at gå med én af dem og får det forhåbentligt til at virke snart. Vi er i gang med en pilot nu (...) for jeg tror, at det er vores akilleshæl, det der med tests og re-tests.

BØRNS VILKÅR

Når jeg skal udvikle, går jeg stort set efter open source, fordi hvis det skal fungere i virkeligheden, og det er betalte leverandører hele vejen rundt, så kan det hurtigt blive dyrt. Så for sådan en praktiker som mig, der gerne vil lege med det og teste ting af, handler det ofte om at dykke ned i en masse dokumentation.

SDU

Det er typisk testfasen og/eller adgangen til open source, der kan lægge en dæmper på virksomhedens anvendelse af taleteknologi, særligt i en dansk kontekst. I takt med at markedet udvikler sig, bliver nuværende udfordringer tydelige. Det afholder dog ikke virksomhederne fra at se muligheder.

Erfaringer med taleteknologi

Interesse i og forventning til taleteknologi

Flere virksomheder ser stort potentiale i taleteknologi, men afventer på nuværende tidspunkt udviklingen lidt endnu.

Vi ved, at de store virksomheder arbejder på det, så der er ikke det samme behov for at kaste vores egne ressourcer efter det, før det bliver vigtigere for os.

NRGi

Store tekniske leverandører tilpasser og udvikler løbende deres modeller, services og løsninger. Det handler for mange om at finde en balance mellem god servicering og digital suverænitæt, herunder intern udvikling samt afdækning af markedet.

Det er jo ikke noget, vi bruger i dag, men jeg tænker, det kommer på et tidspunkt. Det er der ingen tvivl om. Det her kan godt være sådan lidt en konservativ branche, så der går lige lidt tid, inden vi fuldstændig omstiller på sådan noget.

N1

Ud over en digital omstillingsparathed kan andre faktorer som datatilsyn og lovgivning på området være afgørende for, hvordan taleteknologi kan implementeres på arbejdspladsen.

Der skal selvfølgelig være noget governance omkring det, fordi vi er udmærket klar over AI Act og tager det meget alvorligt. Og vi prøver lige nu at etablere et system, så hver gang du tænker AI, så skal du lige spørge ham derovre.

BØRNS VILKÅR

"På dansk halter det lidt efter"

Virksomhederne arbejder alle sammen, i forskellig grad, med taleteknologi på dansk. Flere nævner, at deres eksisterende modellers performance ikke er i top, men brugbar.

Altså, de er jo gode nok til formålet, om der misses et ord hist og pist, det betyder ikke så meget i den kontekst.

BØRNS VILKÅR

Helt konkret viser evalueringerne [på projekterne], at de ikke er gode nok til at forstå, især nydanske elevers accent.

MOEVE

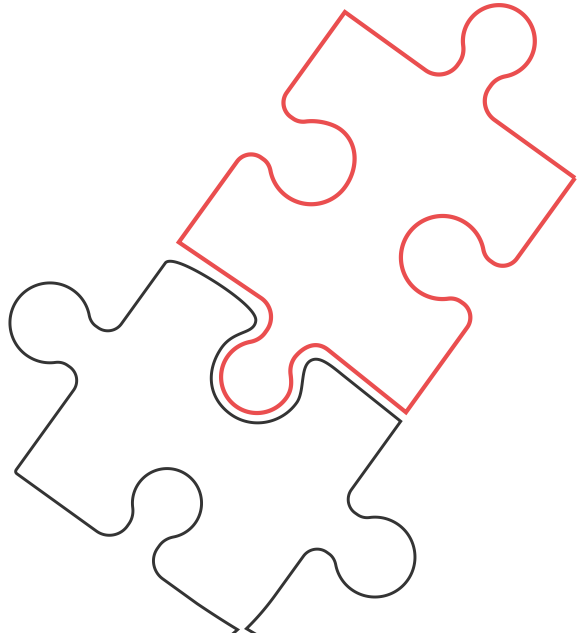
Taleteknologi fungerer fint i langt de fleste sammenhænge, men kvaliteten kan godt falde, når sproget bliver mere komplekst, dialekter og accenter er udprægede, eller der skal arbejdes med specifikt fagsprog. Nogle virksomheder har oplevet, at det særligt kan være en udfordring, når teknologier kombineres.

Mit arbejde handler ofte om en integration af en hel masse forskellige teknologier på tværs, og der er det jo mest optimalt på engelsk eller andre store sprog, og på dansk halter det lidt efter.

SDU

Dette skyldes blandt andet et manglende datagrundlag for testning og udvikling af modeller. Andre faktorer, der har en indflydelse på kvaliteten, er lydniveau, baggrundsstøj, køn m.m. Samtidig er der kommet en større opmærksomhed på bias.

03 CoRals relevans, bidrag og læringer



Virksomhederne er overvejende begejstrede for CoRal projektet, hvor nogle har anvendt projektets ressourcer, og andre følger med fra sidelinjen.

Vi oplever generelt stor interesse for både CoRals datasæt og RØST-modellerne, og flere virksomheder nævner, at de ønsker at kigge nærmere på det snarest muligt. Blandt andet fremhæver Vitec MV, at de er i gang med at undersøge, om CoRals ressourcer har relevans for udviklingen af deres teknologiske løsninger.

Det er fedt med sådan nogle tiltag, fordi det kræver dataindsamling, når man skal lave et eller andet, og vi døjer med det samme, altså inden for tekst ville vi ønske, at der var mere åben datadeling.

VITEC MV

Det er en unik ressource, som jeg tror, at f.eks. Microsoft måske ikke ligefrem kan ramme på kvaliteten, bare for det danske sprog isoleret set.

NRGi

Projektet udfylder dermed et ikke ubetydeligt hul i det danske taleteknologiske landskab ved at stille åbne, mangfoldige og målrettede ressourcer til rådighed, der ellers kan være vanskelige at finde eller opbygge alene.

I takt med at kendskabet til CoRals datasæt og/eller RØST-modellerne vokser, oplever vi overordnet set positiv feedback, hvor virksomhederne især peger på projektets potentiale for mulige anvendelsesscenarier. Det fungerer efterhånden som et konkret alternativ til eksisterende, ofte engelsksprogede services.

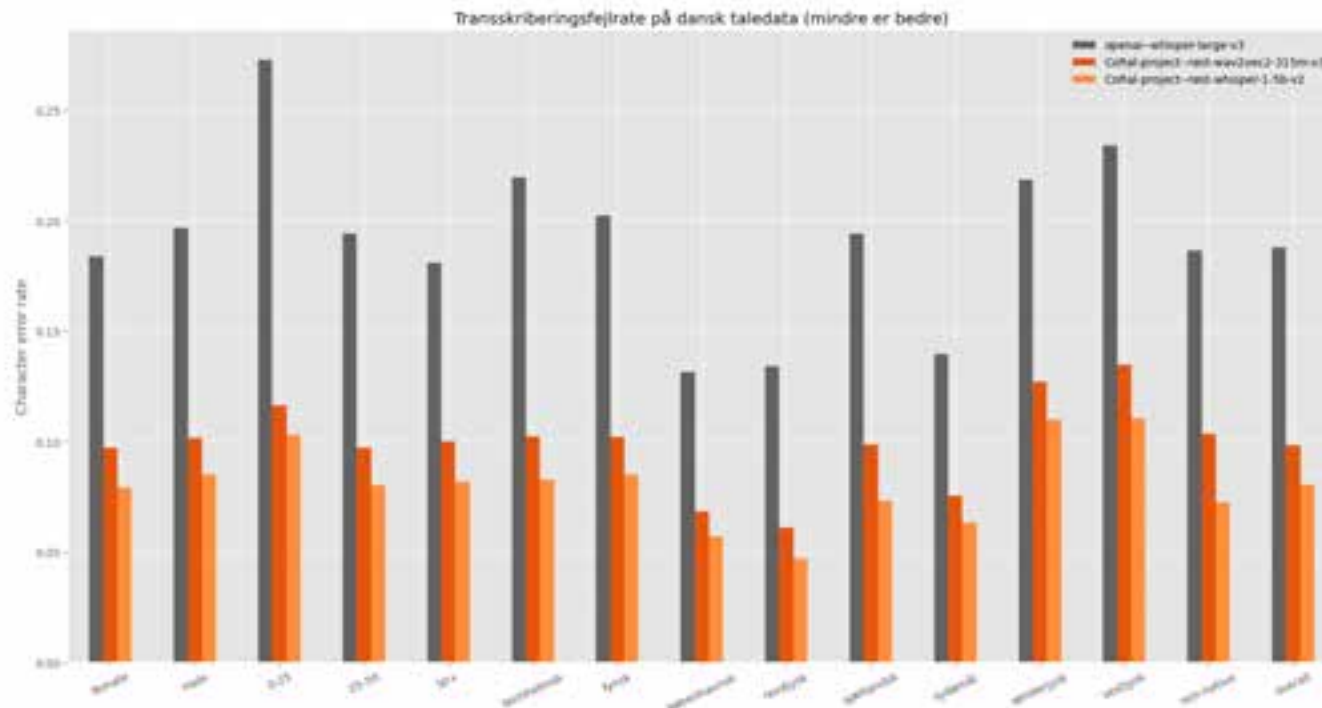
Så er det pludselig en mulighed, at det hele ikke skal være på engelsk, men at man også kan lave projekter eller produkter, hvor man kan bruge dansk.

SDU

Projektets omfang og model performance

CoRal er et af de mest ambitiøse og vellykkede taleteknologiske projekter i Danmark nogensinde. Dette skyldes blandt andet de mere end 1.000 personer, der har lagt stemme til projektet, og ikke mindst deres brede demografiske og demokratiske fordeling. Datasættet og RØST-modellerne kan dog være udfordrede på enkelte dialektområder, da vi har oplevet mindre deltagelse i nogle dele af landet, f.eks. omkring sydøst dialekten, mens rekrutteringen har været særligt stærk i de større danske byer.

Alligevel performer RØST-modellerne bedre end stort set alle tidligere modeller på samtlige grupperinger inden for køn, alder, dialekt samt accent. Gennem hele projektet har vores mål været parallelt at indsamle data og udvikle modeller, og netop denne vekselvirkning har nu resulteret i en bred vifte af værdi-skabende, frit tilgængelige ressourcer. Samlet set er vi med CoRal kommet ét skridt tættere på den teknologiske ligestilling – så dansk taleteknologi en skønne dag kan forstå os alle.



Erfaringer med CoRals ressourcer

Unik ressource

I CoRal er der både et dansk taledatasæt, talesyntese- samt talegenkendelsesmodeller, kaldet RØST, og flere virksomheder fremhæver projektets bidrag på det område.

Jeg ser en stor værdi og et kæmpe potentiale i det, fordi uanset hvordan man vender og drejer det, så har vi mange dialekter og accenter i Danmark. Så det betyder jo, at forskellige løsninger kan blive endnu mere udbredt på dansk.
N1

Det er især mængden af dansk taledata, herunder mangfoldighed og diversitet i sproget, som åbner op for flere og nye muligheder for udvikling af taleteknologi.

Mere diversitet og alle de her ting, som CoRal prøver at hjælpe med ved at give et større datasæt, mere til at træne med, og det er bare fantastisk.

MEDIACATCH

Større dataomfang

Nogle af virksomhederne er begyndt at bruge CoRal datasættet som del af deres træningskorpus, herunder Mediacatch, der blandt andet oplever en forbedring på dialekter.

Mens RØST-modellen vist kæmpede noget mere med for eksempel data talt ud fra en radio- eller tv-udsendelse, end vores egen model gjorde. Der kan det jo gå lidt stærk frem og tilbage.
MEDIACATCH

Flere virksomheder anvender leverandørers 'hyldevare teknologier', hvilket betyder, at datasættet og/eller modellerne skal kunne inkorporeres i eksisterende løsninger. Dette synes at være muligt i de fleste tilfælde.

Vi har talt med vores leverandører, og hvis deres [data] ikke er god nok til at forstå dansk, så har vi CoRal. Og de siger, at man nemt kan plugge det ind.

BØRNS VILKÅR

"Nice and clean"

CoRal datasættet indeholder både oplæst tale og samtaler med fokus på diversitet og demokrati såvel som lyd kvalitet.

Det er nice and clean, hvor jeg før har brugt en masse tid på at rense data for støj eller justeret lyden, har jeg her ikke oplevet problemer med dataen.

NVIDIA, oversat

I tidligere afdækninger har vi dog oplevet, at lyd kvaliteten i datasættet kan være *for god* for virksomhedernes use cases. Blandt andet fordi mange arbejder med virkelighedsnær data (dårlig lyd, støj, spontan tale m.m.) og i stigende grad (fag)specifikke behov.

Nu har vi med ordblinde at gøre og meget af det tekst, der findes tilgængeligt, er korrekt skrevet. Vi kunne godt tænkte os en masse ukorrekt skrevet tekst, og det kan være svært at opsøge.

VITEC MV

Feedback

Til de forskellige virksomhedsdækninger, den foreliggende såvel som tidligere, har vi løbende spurgt ind til feedback, ris og ros.

På tværs af virksomheder og brancher oplever vi fortsat stor tilfredshed med projektets ressourcer, men vi har også modtaget meget specifikke efterspørgsler.

Hvad virksomheder efterspørger

CoRal projektet har haft fokus på at producere et højkvalitets datasæt og tilhørende modeller i forhold til diversitet i det danske sprog. Alligevel kan der være barrierer for at få CoRals ressourcer i anvendelse, og nogle virksomheder fortæller, hvordan det kan skyldes datasættets indhold og/eller interne prioriteringer.

Vi har smalle, målrettede behov, altså botanisk latin, oversættelse dansk-tysk eller nydanske elever, der skal lære et særligt fagsprog inden for SOSU-området.

MOEVE

[Vores løsning] hører forkert, måske endda temmelig ofte, men i det store og hele fanges ting stadig. Så der blev min tid omprioriteret, og jeg kom væk fra CoRal, fordi jeg tænkte, okay, så må jeg genbesøge det, når det bliver mere relevant.

NRGi

Selvom CoRal dækker et bredt spektrum af kvalitet, demografi og sproglig variation, nævner virksomhederne, at deres behov er meget specifikke og målrettede. Dette tyder på, at der er behov for endnu mere open source-data og -modeller, herunder et øget fokus på branchespecifik terminologi og flere færdige plugin løsninger.

"Where the fun begins" – samtaledata

De første releases af CoRals datasæt indeholdt oplæst data, og siden er der kommet samtaledata til. En gennemgående tilbagemelding fra virksomhederne er, at de efterspørger (mere) samtaledata. Det skyldes, at mange use cases er transskribering, herunder talegenkendelse til møder, kundedialog, referatskrivning m.m., hvor der typisk er baggrundsstøj, dårlig lyd og spontan tale med ufuldstændige sætninger.

F.eks. fortæller Nvidia, der har anvendt den første udgave af CoRal datasættet til træningen af deres seneste ASR-model, at kvaliteten og kvantiteten gør, det er nemt at anvende, mens samtaledata virkelig kunne have gjort en forskel. Det er netop samtaler, der kan forbedre taleteknologi, særligt under modeltræning og -udvikling.

Det er her, hvor det bliver sjovt at arbejde med det.

NVIDIA, oversat

Udviklingspotentialer

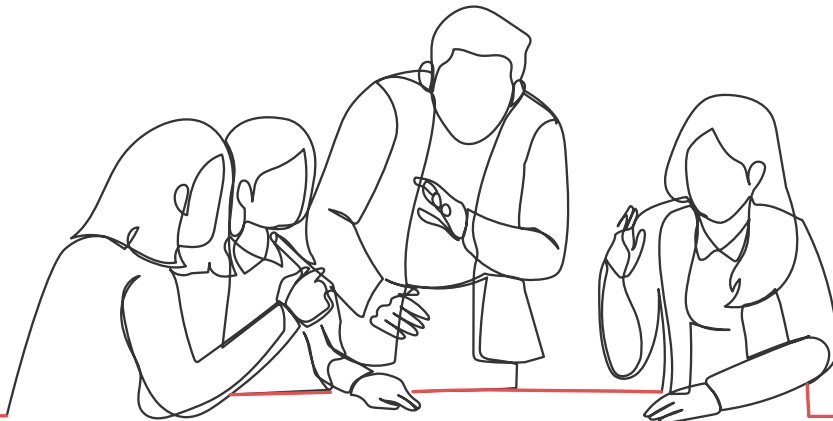
Mere dataindsamling, ønsker og potentialer

Bedst case scenario for forskning og udvikling af taleteknologi er samtaler mellem mennesker, der kender hinanden godt, dvs. hvor tempoet skifter, sætninger afbrydes, og sproget udfoldes.

God følelsesbetonet dialog, f.eks. en diskussion mellem venner, der planlægger en rejse eller et mere teknisk arbejdsmøde. Den type data, hvor ideer udveksles, hvor man kan høre forvirring, livlige diskussioner simpelthen, og ideelt set optaget med flere mikrofoner og forskellige lydkanaler.

NVIDIA, oversat

Flere virksomheder nævner, at det især er hverdagssituationer, herunder data med baggrundsstøj, rumklang og skrattet lyd, som de mangler for at kunne optimere deres modeller, services og løsninger. Det er gældende på tværs af brancher og siloer.



Inputs til at styrke udviklingen

Virksomhederne har mange gode ideer til, hvordan (fremtidig) dataindsamling kan ramme diverse behov til udviklingen af taleteknologi. Der nævnes blandt andet alternative datakilder.

Men det ville helt sikkert være en kæmpe gevinst på kvalitet og diversitet, mere footage fra tv, radio og den slags.

MEDIACATCH

Hvad med film og sådan noget, kunne man f.eks. tage scriptet af en film og køre det igennem?

N1

Dette kan være et alternativ til at få den virkelighedsnære data, som mange virksomheder efterspørger, dog skal man huske at holde øje med den gældende lovgivning på området.

Vi har samtidig oplevet konkret interesse for mere målrettet dataindsamling, herunder nichebegreber, fagsprog og branchespecifik terminologi, som hidtil typisk udfordrer de eksisterende taleteknologiske løsninger og/eller modeller.

Vi ser en øget brug af [løsningen] i sundhedssektoren, specielt fra kommunernes side af, og der er nogle af de her tekniske formuleringer, der godt kan give problemer. Sådan specifikke nicher, fagtekniske nicher, der kunne det være interessant at have noget i en højere kvalitet.

VITEC MV

Innovationskraft i CoRal projektet

En gennemgående tilbagemelding fra virksomhederne er, at både projektets dataindsamling og vekselvirkningen mellem forskning, validering og anvendelse i modeltræningen skaber gode betingelser for udviklingen af dansk taleteknologi.

Det aspekt, som jeg kalder citizen science, altså hvor I netop involverer alle de her gode borgere på etisk forsvarlig vis. Det synes jeg er super spændende og relevant for at få forskning tættere på vores daglige liv, og omvendt.

MOEVE

Vi er jo et lille land, så altså, vi kan ikke regne med at stå først i køen, når der bliver lavet nye versioner af store sprogmodeller, så hvis vi på denne måde kan få vores eget koblet på, er det jo super fint. Det synes jeg er en vigtig del.

BØRNS VILKÅR

I takt med at det danske marked udvikler sig til at være førende i en skandinavisk sammenhæng, er situationen anderledes på f.eks. Færøerne eller Grønland. Det fortæller Vitec MV, hvis tjenester også dækker denne del af Norden.

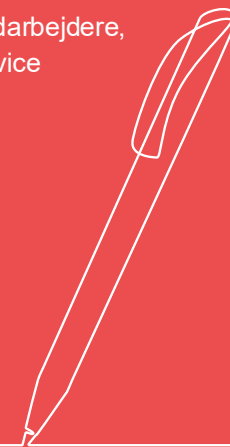
Så der er kvaliteten stadig et issue, og også adgang til ressourcer, der faktisk kan bruge og hoste det.

VITEC MV

Dette understreger værdien i brede, nationale initiativer som CoRal, hvor open source-data og -modeller løfter innovationen, demokratiserer services og udbreder eksisterende biases.

CORALS RESSOURCER KAN ANVENDES I EN LANG RÆKKE TALETEKNOLOGIER, DER BLANDT ANDET KAN HJÆLPE MED OG AFLASTE

- Stemmestyrede hjælpemidler til f.eks. ordblinde, ældre, borgere med synshandicap eller særlige behov
- Diktat af sundhedsjournaler eller dokumentation af samtaler med borgere, kunder og/eller kollegaer
- Voice- og chatbots til f.eks. it-hjælp som nulstilling af passwords, navigations på hjemmesider etc.
- Maskinoversættelse mellem dansk og andre sprog
- Bedre digital oplæsning i f.eks. offentlig transport
- Beslutningsstøtte med relevant viden til medarbejdere, herunder rådgivningsassistenter i kundeservice
- Og mange flere scenarier



04 Afslutningsvis



Hele Danmarks tale-til-tekst model

CoRal projektet har siden begyndelsen af 2023 arbejdet målrettet på at indsamle og udvikle et omfattende taledatasæt samt talesyntese- og talegenkendelses-modeller, der afspejler variationen i det danske sprog. Målet er fortsat at styrke udviklingen af dansk tale teknologi og bringe den op på et niveau, hvor den skal kunne forstå alle borgere – på tværs af dialekter, accenter, alder og køn.

Projektets sounding board, virksomhedsafdækninger samt de mange eksempler på anvendelse, vi har set og oplevet i løbet af projektføreløbet, afspejler faktiske problemstillinger i markedet og peger samlet set på et reelt behov for (endnu) flere open source-datasæt, -løsninger og -modeller.

CoRals ressourcer danner dermed et vigtigt fundament for fremtidig modeltræning og videreudvikling af taledatasæt, men projektet kan ikke alene løse alle udfordringer på området. Vi opfordrer derfor til frigivelse af egne datasæt og/eller modeller, så vi sammen kan skabe et mere innovativt teknologisk økosystem.

FAKTA OM PROJEKTET

TITEL
*Danish Conversational
and Read-aloud speech
dataset (CoRal)*

VARIGHED
2 år og 10 måneder

INNOVATIONSFONDENS
INVESTERING
14.217.380 mio. kr.

SAMLET BUDGET
22.172.400 mio. kr.

Læs mere om
projektet, partnere og
CoRals ressourcer på
[Alexandra Institutet](#).

Har du eller din
virksomhed lyst til at
bidrage med viden,
perspektiver og idéer
til udviklingen af det
tale teknologiske
landskab i Danmark,
kan du tilmelde dig
[TaleTek](#)

Kontakt

Sif Bernstorff Lehmann

Seniorantropolog

Alexandra Instituttet

+45 60 18 58 69

Sif.lehmann@alexandra.dk



Alexandra Institutet hjælper offentlige og private virksomheder med at bruge den nyeste it-forskning og teknologi til at skabe helt nye løsninger. Vores mission er at skabe vækst og velfærd i Danmark.

Sammen kommer vi **#forandigitalt**