



# Taleteknologi med kunstig intelligens

Taleteknologi er systemer som kan tolke hva mennesker sier. Teknologien gjør maskiner i stand til å forstå og reagere på muntlige beskjeder, og å snakke selv.

Utstyrt med kunstig intelligens blir taleteknologien nå raskt bedre på flere områder:

- **Talegjenkjenning.** Å gjøre tale om til tekst og sette sammen ord og setninger.
- **Taletolking.** Å forstå hva folk mener og vil med det de sier.
- **Talesyntese.** Å snakke selv, ved å gjøre om tekst til lyd.
- **Stemmeanalyse.** Å finne unike mønstre i stemmen til en enkeltperson.

Teknologien vil bli en del av de fleste digitale systemer. Det kan styrke språkene i Norge, øke produktiviteten og bidra til at flere får tilgang til arbeid og tjenester. Da må teknologien forstå norske språk og dialekter. Dette vil kreve at Norge satser på teknologien.

## Derfor blir det viktig

### Kan gi arbeid og velferd til flere

[250 000](#) nordmenn har dysleksi, [1 million](#) er pensjonister og [1 av 2](#) uten arbeid har lesevansker. Ved å kunne lese opp og skrive ned tekst, utføre kommandoer og styre ting, som [ovner](#), PC-er, [bilutstyr](#) og velferdsteknologi, kan taleteknologi styrke universell utforming og øke tilgangen til arbeid og informasjon.

### Skjebnetid for norske språk

Norge har flere offisielle skriftspråk og et bredt dialektmangfold. Med taleteknologi som forstår dem kan språksamfunn enklere bevare og uttrykke identitet og kultur i digitale flater. Taleteknologi er én type språkteknologi – systemer som kan tolke både tale og tekst.

## SAMMENDRAG

- » Ny taleteknologi gjør maskiner i stand til å forstå menneskelig tale bedre enn før.
- » Teknologien åpner for at flere kan jobbe og delta i samfunnet, og for økt produktivitet.
- » Å utvikle taleteknologi som forstår norske språk er avgjørende for at språkene skal overleve som bruksspråk i den digitale tidsalder.
- » Tilgang på norske taleopptak, datakraft og gode rammer for norske utviklere blir viktig.
- » Misbruk av stemmeanalyse og kunstig tale byr på utfordringer for sikkerhet og personvern.

I [Digital Agenda](#) beskrives norsk språkteknologi som avgjørende for at norske språk skal overleve som bruksspråk i moderne samfunn, og for demokratisk deltakelse for samisktalende. Alternativt vil engelskspråklig teknologi bli dominerende.

### Spare tid og ressurser

Å diktere går [3–4 ganger](#) raskere enn å skrive. Ved å transkribere møter, avhør, innhold i [pasientjournaler](#) og taler, som på [Stortinget](#), kan taleteknologi derfor øke produktiviteten. Den kan også redusere tid brukt på opplæring. Allerede finnes [fabrikkmaskiner](#) med taleteknologi som kan veilede sine egne operatører.

### Avdekke ny kunnskap

Stemmen er et kroppslig [biometrisk](#) kjennetegn og unik for hver person, som et ansikt. Ved hjelp av mas-

kinlæring kan taleteknologi oppdage unike mønstre i en stemme. Stemmen kan verifisere identiteten til den som snakker og blant annet brukes som [passord](#). Den kan også avdekke informasjon. Forskere eksperimenterer med stemmeanalyse for å screene for [Co-vid-19](#) og for å oppdage [Alzheimer](#) og [Parkinson](#).

### Behovet for norsk

Dersom Norge skal dra nytte av taleteknologiens muligheter er det viktig at teknologien forstår språk som snakkes i Norge. Handlingsplanen [Bærekraft og like muligheter](#) beskriver språkteknologi som en forutsetning for universelt utformede tjenester, og at alle tjenester må fungere på bokmål, nynorsk, samiske språk og dialekter. Det handler om lik tilgang til digitale tjenester, kultur, språk og det å kunne ytre seg.

I dag er engelsk taleteknologi mest utbredt. Teknologien drives framover av tilgang på data, noe som favoriserer store språk. Markedet domineres av selskaper som [Meta](#), [Google](#), [Microsoft](#) og [Amazon](#). De kontrollerer i hvilke flater teknologien kan brukes og hvilke tale- og skriftspråk den skal være tilgjengelig på. Enkelte utenlandske aktører vil kunne tilby noen løsninger som forstår norsk, men kvaliteten vil trolig ikke kunne måle seg med den engelske versjonen.

Dette er kritisk for bokmål og en eksistensiell risiko for nynorsk, dialekter og samiske språk. Fortsetter tjenester å fungere dårlig på norsk kan norsktalende bli ekskludert eller bytte til engelsk. Trenden kan gjøre det krevende for små språksamfunn å beholde språkbrukere, og det kan også true talemangfoldet dersom teknologien kun lærer seg bokmål.

### En nasjonal satsing

Motvekten kan være en nasjonal teknologisatsing som gjør det enklere å utvikle og kommersialisere taleteknologi som forstår norske språk. Da trengs gode vekstforhold for en norsk næring. Det blir også behov for større tilgang på taledata, datakraft og maskinlæring som teknologien kan utvikles med.

### Datatilgang for flere

Å utvikle taleteknologi krever tilgang på taleopptak. De må komme fra omgivelsene der teknologien skal brukes og fra ulike talere. Til enkelte opptak må transkribering følge med, og opptakene må fornyes når språket endrer seg. Teknologien avhenger også av underliggende språkmodeller for å prosessere tekst.

Hovedsatsingen for norsk språkteknologi er [Språkbanken](#) ved Nasjonalbiblioteket. Språkbanken forvalter og gjør delte tale- og tekstdata tilgjengelige for forskere og utviklere. [Divvun](#) og [Giellatekno](#) utvikler

språkteknologi for samiske språk.

I dag distribuerer Språkbanken mindre taleopptak enn det som trengs. Opphavsretten gjør deling utfordrende. [Forskrift til åndsverkloven](#) åpner for deling av pliktavleverte språkdata, som taleopptak fra NRK, men bare til forskning. Dette betyr at verken offentlige eller kommersielle utviklere får tilgang. Dette beskytter åndsverk, men begrenser innovasjon.

Det er også en utfordring å sikre samtykker fra dem stemmene tilhører. En stemme er en personopplysning, og eieren må samtykke dersom den skal brukes til teknologiutvikling. Samtykket må gjentas hver gang taleopptaket skal brukes til noe nytt. [Personvernforordningen](#) (GDPR) artikkel 9 punkt 2d åpner imidlertid for gjentagende behandling av samme personopplysning for enkelte politiske formål. [Finland](#) bruker hjemmelen til å utvikle taleteknologi.

### Skape nye taledata

Det er også behov for helt nye taleopptak, om temaer og fra talere som er underrepresenterte i dag. Taleopptak fra [barn](#), språklige minoriteter og eldre finnes det lite av i Språkbanken. Det samme gjelder uformell prat, som behøves i treningsgrunnlaget dersom teknologien skal fungere i alle situasjoner.

I Finland har Den finske språkbanken, allmennkringkasteren Yle, Det finske klimafondet, universiteter og selskaper [utviklet](#) en nettdugnad for å samle inn hverdagsprat fra folk. Gjennom [nett](#) og [app](#) har det blitt donert totalt 4000 timer prat til nå. Grundige juridiske [vurderinger](#) av personvern ligger til grunn.

Fra enkelte dialektsamfunn, språklige minoriteter og personer med avvikende tale kan det bli ekstra krevende å få samlet inn og gjort nok taleopptak tilgjengelig. Her kan [syntetiske data](#), kunstige taleopptak, muligens spille en rolle. Aktører som Språkbanken bør vurdere hvilke typer taledatasett det er mest nyttig at de prioriterer å fremstille.

### Drive frem et marked

For å øke spredningen av norsk taleteknologi trengs et attraktivt marked for norske utviklere. Markedet må bringe frem løsninger som forstår flere målformer og språk, og motvirke engelskdominans.

I dag er markedet lite og det finnes få stimuleringsordninger for utviklere. Etterspørselen etter teknologien kan derimot være i ferd med å øke. Snart trer EUs [tilgjengelighetsdirektiv](#) og [webdirektiv](#) i kraft i Norge. Disse stiller nye krav til universell utforming av digitale tjenester, som til synstolking av visuelt



innhold på nett. Her kan taleteknologi bli nyttig.

For å anskaffe taleteknologi som oppfyller kravene vil det offentlige trenge solid innkjøpskompetanse. Taleteknologiske løsninger må utvikles spesifikt for funksjonene de skal ha og omgivelsene de skal brukes i for å fungere. Det er derfor viktig at anbudene blir detaljerte og nyanserte.

Offentlige innkjøpere har stor markedsrett. Anbudene kan for eksempel spesifisere hvilke målformer og språk teknologien skal fungere på. Siden flerspråklige løsninger er dyrere og vanskeligere å utvikle bør kartlegginger av hva leverandørene realistisk kan tilby legges til grunn for anbudene.

Det er et politisk spørsmål å vurdere om taleopptak samlet inn gjennom offentlig anskaffede løsninger skal anses som en offentlig ressurs. I så fall kan det vurderes krav til tilbyderne om å dele ikke-sensitive taledata fra løsningene med Språkbanken. Av konkurransehensyn vil enkelte kunne avstå fra å delta i anbud dersom dette blir et krav. En mellomløsning kan være å sette en karantenetid på dataene før de må deles.

Krav om åpen deling kan også vurderes for tale- og språkteknologiske modeller utviklet med pliktavlevert åndsverk. Åpne modeller kan gi flere tilgang til å utvikle teknologi, men også styrke konkurransefortrinnet til de største selskapene. Å dele og holde store modeller oppe og oppdaterte krever mye datakraft og kompetanse. Det har selskaper som Google mest av i dag. Norske utviklere vil trenge solid tilgang på infrastruktur for å lage, lagre og bruke slike modeller dersom de skal klare å konkurrere mot de største.

### Sikre datakraft og ekspertise

Å utvikle taleteknologi krever mye datakraft. Det er fordi regneoperasjonene som må til for å trene modellene er komplekse. Teknologien er komplisert å utvikle, og fordrer kunnskap om både lingvistik, datavitenskap og elektronisk systemdesign.

Tilgang til Norges [tungregningsressurser](#) - de kraftigste datamaskinene - styres av universitetene og det statlige selskapet Sigma 2. Disse bestemmer hva ressursene skal gå til. Tilgang må søkes om og tildeles etter [kriterier](#) som nyskaping og forskningsverdi. Ressursene er ikke øremerket ulike temaområder.

Både [NTNU](#), [Universitetet i Oslo](#) og [Sigma 2](#) bygger nå ut regneressursene sine. Framover blir det en viktig politisk oppgave å sikre at forskning og teknologiutvikling av særlig interesse for samfunnet får god tilgang til datakraft og avansert maskinlæring.

Det blir også avgjørende at forskere og utviklere blir satt i stand til å bruke ressursene og forvalte data forsvarlig og får god tilgang på juridisk rådgivning.

Utvikling av norsk taleteknologi forutsetter tilgang på personer med relevant utdanning og kompetanse. I [Langtidsplan](#) for forskning og høyere utdanning er muliggjørende og industrielle teknologier et av temaområdene. Innenfor temaområdene er det en politisk beslutning hvilke teknologier som skal prioriteres. Tale- og språkteknologi bør være en sterk kandidat.

### Støtte små språksamfunn

Teknologi må kunne tas i bruk for å bli nyttig. I dag unngår ofte de største leverandørene av taleteknologi å gjøre grensesnittene til programmene og de digitale assistentene sine kompatible med teknologi andre har utviklet. Dette går særlig utover små språksamfunn, som forhindres fra å ta i bruk egenutviklet teknologi.

Slik risikerer små språksamfunn som det samiske å miste kontroll over i hvilke flater og med hva slags digital teknologi språkene kan brukes. Det oppstår en situasjon der språksamfunnenes egenutviklede løsninger kan være bedre enn dem som de største leverandørene tilbyr, men uten at de får brukt dem.

Her kan regjeringen og Sametinget gjennom et felles nordisk initiativ gå i dialog med de største leverandørene og stille krav om kompatibilitet med språksamfunns egenutviklede teknologi.

### Motvirke misbruk

Samtidig som det trengs en satsing, blir det viktig å unngå at informasjonen i en stemme misbrukes og at teknologi blir brukt til svindel eller desinformasjon.

### Emosjonell overvåking

Stemmeanalyse kan brukes til å identifisere folk, men også til å oppdage sinnsstemninger og følelser. Forskerne er [uenige](#) om hvor nøyaktig slik [bruk](#) av teknologien er. Likevel blir stemmeanalyse brukt til å anslå folks [alder](#), [følelser](#), [kjønn](#) og [høyde](#).

For eksempel kan Amazon Alexa høre om noen er forkjølet og automatisk bestille [hostemedisin](#). Spotify fikk i 2021 [godkjent](#) et [patent](#) for musikkbefalinger basert på stemmeanalyse, til stor protest fra [musikkbransjen](#) og en rekke ideelle [organisasjoner](#). Systemet skal kunne sortere og kategorisere lytterne basert på sinnsstemning, kjønn, alder og dialekt.

Det er ikke forbudt med slike analyser. Personvernforordningen [tillater](#) behandling av biometriske data så lenge eieren samtykker. I EU-kommisjonens



[forslag](#) til regulering av kunstig intelligens, Artificial Intelligence Act, stilles det krav om at personer som eksponeres for følelsesgjenkjenning skal få vite om det. Forslaget stiller også [strengt krav](#) til utvikling og bruk av slike systemer innen områder som helse, arbeidsliv, rettsvesenet, utdanning og velferd.

Det er et politisk spørsmål å vurdere om bruk av stemmeanalyse skal være tillatt i Norge og hva det kan brukes til. Uansett blir det viktig å sikre at beslutninger basert på stemmeanalyse blir mulige å forklare og forstå for dem som utsettes for dem.

### Stemmekloning og desinformasjon

Avansert maskinlæring gjør det stadig enklere å [klone](#) og [gjenskape](#) en stemme, og å [generere](#) nye taleklipp med den. Dette reiser flere etiske og politiske spørsmål rundt hvem som eier en stemme, tillit til digitalt innhold og menneskeverd i møte med maskiner.

Det skapte [debatt](#) da en filmskaper [klonet](#) og brukte en død persons stemme i en dokumentar om avdøde. Selskapet [HereAfter AI](#) lager og lar folk snakke med digitale reproduksjoner av [døde personer](#), trent og basert på taleopptak av de avdødes stemmer.

Stemmekloning og kunstige stemmer kan også brukes til å [svindle](#) og lage falskt innhold, såkalte «[deepfakes](#)». I mars 2022 sirkulerte en [video](#) der Ukrainas president Volodymyr Zelenskyj ba ukrainerne overgi seg. Zelenskyj måtte selv gå ut og [si](#) at videoen var falsk.

I forslaget til Artificial Intelligence Act stilles det krav om merking av manipulerte og kunstig fremstilte bilder, lydklipp og videoer. I tillegg skal personer informeres dersom de kommer i kontakt med kunstig intelligente systemer, som prateroboter.

Et tiltak som kan redusere usikkerhet kan være å presisere hvem som eier og har ansvaret for en klonet eller kunstig stemme. Det kan også vurderes å lovfeste en rett til kontakt med mennesker i det offentlige.

### Anbefalinger

Skal taleteknologi bli en berikelse må forutsetningene for datadeling, en næring og forbrukerbeskyttelse være på plass. Norske myndigheter bør derfor:

## ANBEFALINGER

- » Beskrive i [Digitaliseringsrundskrivet](#) hvordan tale-data enklere kan deles for teknologiutvikling og gjøre det til en plikt for forskere og offentlige aktører å dele alle tale- og tekstdata de kan med Språkbanken.
- » Endre Forskrift til åndsverkloven § 4 slik at Nasjonalbiblioteket kan dele opphavsrettsbeskyttet åndsverk til utvikling i offentlig regi og at kommersielle utviklere kan få søke om tilgang til ressursene.
- » Kreve at vinnere av offentlige anbud deler ikke-sensitive taledata som løsningene deres samler inn med Språkbanken etter en satt karantenetid.
- » Finansiere og gjennomføre nasjonale og nordiske nettdugnader for å samle inn hverdagsprat.
- » Vurdere om tale- og språkmodeller utviklet med pliktavlevert åndsverk skal gjøres åpent tilgjengelige.
- » Prioritere forskning og forskningsdrevet innovasjon på språkteknologi som muliggjørende teknologi i Langtidsplan for forskning og høyere utdanning.
- » Utarbeide en veileder for offentlige innkjøp av taleteknologi, som tar i betraktning nye krav i EUs webdirektiv og tilgjengelighetsdirektiv.
- » Kreve at store leverandører gjør tjenester kompatible med språksamfunns egenutviklede teknologi.
- » Arbeide for bedre økonomiske rammevilkår for utviklere av tale- og språkteknologi.
- » Lovfeste en rett til å bli behandlet av mennesker i offentlig sektor, og juridisk avklare eierskapet til og ansvaret for klonede og kunstig genererte stemmer.

Redaksjon: Tore Tennøe / Jonas Engestøl Wettre

Publisert: desember 2022

Kontakt: [post@teknologiradet.no](mailto:post@teknologiradet.no) / [www.teknologiradet.no](http://www.teknologiradet.no)

Ekspertgruppe: Knut Kvale (Telenor Research), Jon Atle Gulla (NTNU og NorwAI), Line Adde (Aidn AS), Per Erik Solberg (Nasjonalbiblioteket), Torbjørn Karl Svendsen (NTNU) og Sjur Nørstebø Moshagen (Divvun og Universitetet i Tromsø)

Teknologirådet gir råd til Stortinget og regjeringen om ny teknologi