

Test af AI-tolk på borgersamtaler i BIF

Kort resumé:	Test af AI-tolkeværktøjer i borgersituationer i takt med den teknologiske og markedsmæssige modenhed til analyse af de juridiske og organisatoriske muligheder for implementering af AI-tolkning i beskæftigelsesindsatsen.
Fremstillende forvaltning:	Beskæftigelses-, Integration og Erhvervsforvaltningen

1. Beskrivelse af forslag (fremgangsmåde og effekt)

AI-tolk, når det er muligt – menneskelig tolk, når det er nødvendigt. Det kan meget vel være visionen for en fremtidig implementering af AI-tolke i borgersamtaler på beskæftigelsesområdet. Beskæftigelses-, Integrations- og Erhvervsforvaltningen har med de seneste års markante indvandring, som følge af krigen i Ukraine, oplevet markant stigende udgifter og administrative byrder i forbindelse med tolkning.

Dertil har forvaltningen ofte svært ved at finde gode og stabile tolke til de mange sprog og dialekter, der mødes i samtaler med kommunens borgere. Det resulterer både i svingende kvalitet og ensartethed i tolkningen og i aflyste samtaler. Det påvirker både sagsbehandlingen, borgerens oplevelse og medarbejdernes mulighed for at vejlede korrekt.

Innovationsforslaget tager derfor afsæt i behovet for at finde en digital, effektiv og tilgængelig form for tolkning, der kan anvendes som erstatning for eller supplement til menneskelige tolke med henblik på at øge serviceniveauet og frigøre ressourcer.

Hvad har vi lært af foranalysen?

Baggrunden for innovationsforslaget er en foranalyse gennemført i 2025, der havde til formål at afdække forretningsbehovet, de juridiske rammer, den teknologiske udvikling og den markedsmæssige modenhed af AI-tolkeløsninger. Dertil foretog forvaltningen en mindre afprøvning af fire forskellige værktøjer under udvikling for at få et øjebliksbillede af kvalitet og brugeroplevelse af løsningerne. Resultaterne følger her:

Forretningsbehovet er tydeligt

Beskæftigelses-, Integrations, og Erhvervsforvaltningen afholdt i 2024 ca. 14.000 samtaler med borgere med tolkebehov. Samlet set brugte forvaltningen over 13 mio. kroner på tolkning mellem borger og sagsbehandler. Hvor det tal varierer naturligt med den årlige migration mod Danmark – særligt efter krigen i Ukraine – bruger forvaltningen hvert år millioner i direkte omkostninger til tolke, ligesom estimerer fra foranalysen viser, at der mindst bliver brugt 50.000 minutter eller over 100 arbejdsdage på administration af tolke. Den nuværende model med manuelle bookinger og fysiske tolkesessioner skaber flaskehalse, forsinkelser og sårbarhed over for sygdom, aflysninger og rekrutteringsproblemer.

Hjemmelsproblematikken bliver løst med beskæftigelsesreformen

Der fandtes ikke i 2025, mens foranalysen stod på, en tilstrækkelig klar lovhjemmel til at anvende AI-baserede tolkeløsninger i borgersamtaler – hverken i sektorlovgivningen eller i forvaltningsloven. Det ændrer sig med den nye beskæftigelsesreform, der træder i kraft i februar 2026, hvor der er udtrykt en klar ambition om øget anvendelse af kunstig intelligens. Det afspejles i den tilhørende AI bekendtgørelse, hvor der er udspecificeret en hjemmel til anvendelse af kunstig intelligens i beskæftigelsesindsatsen som heller ikke er afgrænset i forhold til behandling af personoplysninger. Både forvaltningens egne jurister, leverandører og KL vurderer derfor ud fra en indledende analyse af bekendtgørelsen og med forbehold for, at den stadig er i høring ved forslagets tilblivelse, at vejen er banet for implementering af tolkeværktøjer.

Digitale tolkeløsninger er et felt i udvikling

Tale-til-tekst modeller er en teknologi i hastig udvikling. Således publicerer Microsoft og OpenAI hver måned nye opdateringer både i kvalitet og præcision i deres sprogmodeller, og det er nu så langt, at det har banet vejen for udviklingen af AI-tolkeløsninger.

Under foranalysen er der foretaget afprøvninger af de fire løsninger, der er på vej på det danske marked, og som alle er tidligt i deres udviklingsfase. Fokus for afprøvningerne var dels at få et øjebliksbillede af kvaliteten, og dels at få en bedre fornemmelse for brugeroplevelsen i samtalen. I en minitest over to dage med 32 simulerede borgersamtaler mellem sagsbehandlere på tværs af forvaltningen, stod det først og fremmest klart, at løsningerne stadig er i udvikling. Kvaliteten i transskribering og tolkning varierer, og de forskellige platforme (fx mobil vs. desktop) har hver især fordele og ulemper. De står alle stærkest på de store sprog som russisk og ukrainsk, der udgør 57% af tolkesamtalerne, men svagere på arabisk og særligt på mindre sprog som urdu.

Som Dataetisk Råd dog peger på i en nyligt udgivet rapport om *Generativ AI*, er det dog uundgåeligt, at sprogmodeller vil lave fejl — teknologien vil aldrig blive fejlfri. Den relevante vurdering handler derfor ikke om at opnå nul fejl, men om at fastlægge hvilken type fejl og hvilken mængde af fejl, der er acceptabel i det konkrete brugsscenarie — og at vælge testmiljøer, der matcher den risikoprofil, vi kan håndtere. Hvor konklusionen er, at teknologien endnu ikke er klar, så indeholder hver af løsningerne interessante forskelle i kvalitet og brugeroplevelse, og flere af leverandørerne melder, at de forventer en markedsklar og moden løsning klar allerede i første halvdel af 2026.

Innovationsforslag: Yderligere test er nødvendigt som forløber til investeringsforslag

Foranalysen i 2025 har skabt et solidt grundlag for at forstå det juridiske, tekniske og forretningsmæssige potentiale ved brugen af AI-baserede tolkeløsninger. Analysen har dog også vist, at der fortsat er behov for teknologisk udvikling og empirisk afprøvning i beskæftigelsesindsatsen, før der kan udarbejdes et endeligt investeringsforslag.

Forvaltningen foreslår derfor et innovationsforslag med henblik på løbende at teste et eller flere af de nuværende AI-tolkeværktøjer i markedet i 2026 i takt med, at de bliver meldt klar til implementering. Det inkluderer test på sprog som russisk, ukrainsk og arabisk, som er de mest udbredte sprog blandt ledige borgere med tolkebehov i København, men så vidt muligt også test på sprog som grønlandsk og urdu, hvor det er særligt svært at finde gode tolke.

De løbende nålestikstests skal give praktisk erfaring med teknologiens kvalitet og brugervenlighed, alt imens der udvikles undervisningsmateriale, der kan understøtte en fremtidig implementering og belyse, hvordan brugen af AI-tolke påvirker sagsbehandlere og administrativt personale.

Afprøvningen vil ske med forskellige leverandører, der repræsenterer forskellige tekniske tilgange og spænder over forskellige sprog. Afhængigt af udviklingen af løsningerne og det endegyldige juridiske spillerum i den nye beskæftigelsesreform vil afprøvningsne indledes på lav-risiko situationer som videregivelse af beskeder i uplanlagte samtaler, før de afprøves i mere relevante og planlagte borgersamtaler.

En del af projektet består også i at afklare, hvordan AI-tolke håndterer centrale faglige termer, teknisk sprog og udtryk, der anvendes i beskæftigelses- og integrationsindsatsen. Det omfatter både præcision i ordvalg, forståelse af kontekst og gengivelse af nuancer, som menneskelige tolke normalt håndterer.

Innovationscasen skal derfor identificere, *hvilke sproglige og faglige elementer, der er kritiske for korrekt kommunikation*, og hvor teknologien leverer tilstrækkelig kvalitet. Dette skal danne grundlag for at afgrænse, i hvilke samtaler AI-tolkning kan anvendes forsvarligt, og hvor menneskelig tolkning fortsat er nødvendig.

Samlet skal det danne grundlag for en opdateret vurdering af de økonomiske og administrative gevinster for en investeringscase i 2027 med afsøgning af muligheden for et tværgående forslag.

Projektet forankres i Kontor for Digitalisering (BIF) med deltagelse fra udvalgte centre (CSI, CBA, CBE, m.fl.) og i samarbejde med Koncern IT. Herudover etableres en følgegruppe igennem AI-netværket, hvor de øvrige forvaltninger inddrages i erfaringerne fra projektet og der på sigt kan vurderes om teknologien har et tværgående potentiale.

2. Økonomi

Projektet foreslås gennemført som et étårigt innovationsforløb finansieret af Innovationspuljen. Midlerne skal dække omkostninger til projektledelse, teknisk opsætning og licenser, afklaring af sproglig kvalitet samt inddragelse af medarbejdere og borgere i testfasen.

De estimerede omkostninger til test er baseret på dialog med de leverandører, der var en del af markedsafdækningen i den indledende foranalyse. Det øvrige ressourcebehov er estimeret ud fra timeforbruget under foranalysen, samt erfaringer fra tidligere gennemførte og igangværende innovationscases i BIF.

Innovationsmidlerne dækker de direkte udgifter til projektet, mens de deltagende centre forventes at bidrage med egenfinansieret arbejdstid til test og evaluering.

Projektet organiseres med en projektleder i Kontor for Digitalisering (BIF) og en tværgående styregruppe med repræsentanter fra CSI, CBA, og CBE. Derudover etableres en projektgruppe samt en følgegruppe af repræsentanter fra de øvrige forvaltninger.

Som innovationsprojekt er formålet ikke at skabe varige effektiviseringer i 2026, men at modne et beslutningsgrundlag for en efterfølgende investeringscase i 2027. Hvis testen viser tilfredsstillende resultater, forventes der at kunne udarbejdes en businesscase med konkrete estimater for:

- reduktion i udgifter til tolkning,
- øget tilgængelighed for borgere,
- samt tidsbesparelser i sagsbehandlingen.

BIF vil være det første forvaltningsområde, hvor AI-tolk bliver taget i brug i borgersamtaler i KK. Potentialer er stort også i andre forvaltninger med direkte borgerkontakt som KFF, SOF, BUF og SUF. En del af projektet skal understøtte, hvordan viden fra brug af Altolk på beskæftigelses- og integrationsområdet kan give relevante erfaringer til andre forvaltninger i KK. Med afsæt i det og via dialog med KIT skal det ligeledes undersøges, om AI-tolk på sigt kan blive en fælles it-kapabilitet på tværs af KK.

Projektet afsluttes med en evalueringsrapport og anbefalinger til videre udmøntning, finansiering og implementeringsmodel. Den samlede budgetramme forventes at udgøre 1,0 mio. kr., fordelt på følgende hovedposter:

Tabel 1. Forslagets samlede økonomiske konsekvenser

	1.000 kr. 2026 p/l						
	Styringsområde	2026	2027	2028	2029	2030	2031
Implementeringsomkostninger							
Projektledelse (inkl. evaluering, dokumentation og rapportering)	Service	350					
Teknisk opsætning, licenser og support (leverandører af tolkeløsninger)	Service	400					
Dataetiske og sprogfaglige afklaringer	Service	100					
Afklaring af tekniske forhold i forbindelse med test på borgersituationer.	Service	200					
Samlede implementeringsomkostninger		1.000					

3. Kommende investeringsforslag og kriterier til prioritering

Innovationsprojektet skal danne grundlag for et kommende investeringsforslag i 2027, der skal realisere de økonomiske og organisatoriske gevinster ved at anvende AI-tolkeløsninger på de borgersamtaler, der lovligt og fagligt kan understøttes med AI-tolkning.

Investeringsforslaget forventes at bygge på erfaringer, undersøgelser og data fra testforløbet i 2026, herunder teknologisk kvalitet, juridisk anvendelighed, medarbejder- og borgeroplevelse, samt de estimerede driftsmæssige effektiviseringer.

Det forventes, at erfaringerne fra innovationscasen vil danne grundlag for en gradvis udrulning af AI-tolkeløsninger i takt med udviklingen i teknologien og i markedet – det vil sige i takt med, at kvaliteten i de enkelte sprog forbedres, og at leverandørerne udvikler funktionalitet, som gør teknologien anvendelig i et bredere udsnit af kommunens samtalsituationer – fx møder med baggrundsstøj, samtaler med flere

deltagere eller komplekse rådgivningsforløb. En sådan trinvis implementering vil reducere risici og samtidig sikre, at kommunen udnytter teknologien, hvor den giver størst værdi først.

De estimerede udgifter til et kommende investeringsforslag er den forventede årlige udgift for Beskæftigelses, Integrations og Erhvervsforvaltningen på baggrund af en indledende prissætning fra leverandører. Såfremt innovationscasen munder ud i en tværgående investeringscase for flere forvaltninger, vil udgiftsniveauet selvfølgelig stige.

Vi forventer at indhente support fra KIT for 50-100tkr i arbejdet med den tekniske afklaring.

Tabel 2. Kriterier til prioritering.

Kommende investeringsforslag		Kriterier til prioritering			
Udgift 1.000 kr. 2026 p/l	Forventet vedtagelse	Potentiale for varig effektivisering (styringsområde)	Tværgående effektivisering	Potentiale for storskala	Modne anvendelse af ny teknologi
1.300 kr.	OFS 26/27	4-8 mio. kroner	X	X	X

4. Implementering og opfølgning

Projektet afsluttes med en evalueringsrapport og præsentation for IT-kredsen og styregruppen i december 2026. Rapporten vil sammenfatte testresultater, anbefalinger for juridisk og teknisk anvendelse samt en foreløbig business case. Resultaterne deles med de øvrige forvaltninger gennem følgegruppen og vil indgå som dokumentation i et eventuelt investeringsforslag i 2027.

Tabel 3. Opfølgningsmål

Opfølgningsmål	Hvordan måles opfølgningsmålet?	Hvem er ansvarlig for opfølgning?	Hvornår gennemføres opfølgningen?
Realisering af innovationsforslag			
Teknisk kvalitet og nøjagtighed	Fejlrate (accuracy errors) målt på stikprøver af samtaler, suppleret af medarbejdervurdering.	KfD	Løbende under test og samlet evaluering i Q4 2026
Brugervenlighed og oplevet værdi	Survey til sagsbehandlere og borgere (tilfredshed, tillid, oplevet forståelse).	KfD	Efter hver testperiode (midtvejs og slut)
Juridisk vurdering	Analyse af ny lovgivning samt efterfølgende bekendtgørelse ud fra mulige brugsscenarier	KfD	Q2 2026
Økonomisk potentiale	Sammenligning af tolkeudgifter og tidsforbrug før og under test, samt scenarieestimat for effektivisering. Vil udgøre grundlaget for fremtidig investeringsforslag.	KfD	Q4 2026

Organisatoriske erfaringer og kompetencebehov	Interviews og workshop med medarbejdere om roller, ansvar og læringsbehov.	KfD	Løbende opsamling og afsluttende Q4 2026
Skalerings- og tværgående potentiale	Afsøge interesse og mulige gevinster ved implementering i øvrige forvaltninger med henblik på vurdering af om der skal udarbejdes et tværgående investeringsforslag.	KfD	Q4 2026 – input til investeringscase 2027

5. Risikovurdering

Afprøvningen af AI-baseret tolkning indebærer både juridiske, tekniske og organisatoriske risici. Da der er tale om en innovationscase for et felt under modning – både teknologisk og i mulighederne for overhovedet at anvende kunstig intelligens i borgersamtaler, er det centralt at arbejde aktivt med risikostyring. Projektet følger kommunens retningslinjer for databeskyttelse, informationssikkerhed og ansvarlig brug af kunstig intelligens, og der føres en løbende risikolog, som gennemgås kvartalsvist i styregruppen.

Juridisk risiko

Til februar 2026 træder en ny beskæftigelsesreform i kraft, der indeholder en hjemmel til brug af AI i sagsbehandlingen. Den umiddelbare vurdering af den nye AI bekendtgørelse blandt BIF's egne jurister og KL er, at vejen er banet for anvendelsen af AI-tolke. Uagtet vores egen interne vurdering af det juridiske spillerum, er vi dog i en situation med forhøjet opmærksomhed om anvendelsen af kunstig intelligens. Det kan føre til ankesager og domsafgivelser, som potentielt set vil forme og ændre brugsscenarierne. For at imødegå denne risiko gennemføres testen kun i situationer, hvor der foreligger samtykke, og hvor forvaltningen vurderer lovgrundlaget som tilstrækkeligt. Der kan være en yderligere afgrænsning for borgere, der går på tværs af sektorlove, der ikke har samme hjemmelsgrundlag. Testen vil i så fald ekskludere disse grupper af borgere.

Risikovurdering: Lav

Databeskyttelse og informationssikkerhed

AI-tolkeløsninger håndterer lyd- og tekstdata, som kan indeholde personfølsomme oplysninger. Risikoen for utilsigtet behandling eller lagring af data uden for EU håndteres ved at anvende leverandører, der opfylder kommunens krav til datasikkerhed og ved at sikre, at al behandling sker inden for EU/EØS. Koncern IT indgår databehandleraftaler med leverandørerne og fører tilsyn med hosting, adgangskontrol og sletningsrutiner.

Risikovurdering: Lav

Teknisk kvalitet og præcision i sprog og fagtermer

AI-tolke kan levere upræcise eller fejlagtige oversættelser, særligt i samtaler med baggrundsstøj, dialekter eller flere talere. Dette kan skabe misforståelser eller fejlagtig vejledning. Risikoen begrænses ved at afgrænse testen til lavrisiko-samtaler på de sprog, der udviser mest præcision og kvalitet i oversættelsen, gennemføre kontrolsamtaler med menneskelig tolk som reference og analysere fejltypen systematisk. Derudover vurderes teknologiens evne til at håndtere centrale fagtermer og teknisk sprog, da mangelfuld oversættelse her kan påvirke samtaleforståelse.

Risikovurdering: Middel

Etiske og bias-relaterede risici

Sprogmodeller kan gengive bias, stereotype tolkninger eller forskelsbehandling på baggrund af sproglig eller kulturel variation. Denne risiko imødegås gennem kvalitetskontrol, manuel gennemgang af samtaler og et etisk review baseret på Dataetisk Råds anbefalinger om fejltyper, gennemsigtighed og ansvarlighed.

Risikovurdering: Lav

Brugeraccept og tillid

Medarbejdere og borgere kan opleve utryghed ved at anvende AI som tolk, hvilket kan påvirke anvendelsesgrad og datakvalitet. Projektet imødekommer dette gennem tydelig kommunikation, frivilligt samtykke og inddragelse af medarbejdere i både design og evaluering. Erfaringer fra brugerundersøgelser dokumenteres som en del af evalueringen.

Risikovurdering: Middel

Økonomisk og driftsmæssig risiko

Licenser, API-forbrug eller teknisk opsætning kan vise sig dyrere end forventet, eller leverandørens løsning kan opleve ustabil drift. Risikoen reduceres ved at fastsætte et klart budgetloft, gennemføre leverandørdialog inden teststart og følge økonomien løbende. Der anvendes flere leverandører for at reducere afhængighed af én aktør.

Risikovurdering: Lav-middel

Organisatorisk forankring

Der er risiko for, at projektets resultater ikke implementeres, eller at viden går tabt efter afslutning. For at sikre forankring kobles innovationscasen direkte til arbejdet med et kommende investeringsforslag, og resultaterne forankres direkte i jobcentrene.

Risikovurdering: Lav-middel

Administrative bilag til innovationsforslag

6. Hvem er hørt?

Tabel 4. Godkendelse af Center for Økonomi

	Ja/Nej	Dato for godkendelse
Center for Økonomi, Økonomiforvaltningen	Ja	Godkendt af CFØ 9. januar 2026 Godkendt af CDØ i november 2025

Tabel 5. Inddragede samarbejdspartnere

Relevante samarbejdspartnere i kommunen, der er inddraget
• Innovationsforslaget er sendt i høring hos KIT (dec. 2025) • Øvrige forvaltninger er orienteret, og BIF agter at indkalde forvaltningerne primo 2026 til et indledende orienteringsmøde om en mulig tværgående investeringscase fra 2027.