

Steg-för-steg guide

Transkribering av möten

2025-04-30





Transkriberingstekniker och verktyg

Detta är en steg-för-steg guide som jag tog fram till mig själv för att bättre transkribera möten. Alla transkriberingsverktyg har sina fördelar, inget av verktygen är perfekt på allt. Så välj det som passar dig bäst på den typ av möten du är.

Allt förändrar sig också så snabbt, så detta är hur det ser ut just nu. För ett par månader sen fanns inte KB-Whispers fantastiska svenska version eller funktionaliteten i Word. Så det är bara att hänga med.

Generellt så rekommenderar jag att börja testa med Teams eller Word, därefter kan det bli aktuellt med Klang.ai eller annan specifik tjänst. Sen blir det för de lite nördiga att testa med KB-Whisper. Troligt är att den sistnämnda kommer bli inbyggd i andra lite mer åtkomliga verktyg.

Återkoppla gärna till lyssnare@projektledarpodden.se med vad som kan förbättras eller andra guider ni vill ha.

Sollentuna 2025-04-30

Mattias Ejbe

Använda transkribering i Microsoft Teams (Microsoft 365 Enterprise)

Kort introduktion

Transkribering i Microsoft Teams innebär att tal under möten automatiskt omvandlas till text i realtid. Funktionen fångar upp allt som sägs och visar en löpande direktavskrift under mötet, komplett med varje talares namn och tidsstämpel. Detta gör det enklare att följa med och spara en skriftlig kopia av diskussionen.

Förutsättningar och fakta

- **Licenser:** Transkriberingsfunktionen är tillgänglig för organisationer med Microsoft 365 Enterprise E3/E5 eller Business Standard/Premium-abonnemang. (Den finns inte i privata/gratis Teams-versioner – du behöver en företagslicens för att använda transkribering.) Se även till att IT-administratören har aktiverat transkribering i Teams inställningar, annars kan knappen vara inaktiv.
- **Enheter:** Live-transkribering stöds endast i Teams skrivbordsapp (Windows/Mac). Mobilappen kan för närvarande bara visa transkriptionen efter att mötet avslutats, inte i realtid under mötet. (Teams på webben har begränsat stöd – för bästa resultat, använd skrivbordsappen.)
- **Språk:** Ursprungligen fanns stöd bara för engelska, men idag stöds över 40 olika språk för transkribering, inklusive svenska. Svenska (Sverige) finns bland de tillgängliga språken. Generellt är noggrannheten högst för engelska (framförallt US/UK-engelska) tack vare mer träningsdata, men även svenska transkriberingar fungerar bra under klara förhållanden. Microsofts automatik når cirka 80% exakthet i ideala fall, vilket betyder att det oftast fångar helheten korrekt men ibland kräver smärre justeringar. Otydligt tal, bakgrundsljud eller dialekt kan sänka precisionen något, så tala tydligt för bästa resultat.

Steg-för-steg-instruktion för transkribering

Följ dessa steg för att använda transkribering i ett Teams-möte och sedan ladda ner texten:

1. **Planera mötet:** Schemalägg ett Teams-möte i kalendern (transkribering fungerar endast i schemalagda möten, inklusive kanal-möten – inte i spontana "Meet now"-möten eller samtal).
2. **Starta mötet:** Gå med i mötet som organisatör eller presentatör.
3. **Påbörja transkribering:** Klicka på Fler åtgärder (de tre prickarna ... i möteskontrollerna). Välj sedan "Spela in och transkribera" och klicka på "Starta transkription". (Om du väljer att spela in mötet så startar transkriberingen automatiskt samtidigt.)

4. **Bekräfta språket:** Teams kan bekräfta vilket språk som talas i mötet första gången du startar transkriberingen. Se till att välja rätt talat språk (t.ex. Svenska) i prompten och klicka Bekräfta. (Tips: Om mötets språk ändras eller identifieras fel kan organisatören uppdatera språk via Språkinställningar i transkriptionsfönstret under mötet.)
5. **Under mötet:** När transkriberingen är igång visas texten live på skärmen för alla deltagare. Teams visar vem som pratar (namn) före varje uttal samt tidsangivelse. Deltagare får även en avisering om att mötet transkriberas och kan välja att dölja sitt namn i avskriften vid behov. Obscent språk filtreras automatiskt och ersätts med *** i transkriptionen.
6. **Stoppa transkriberingen:** Du kan när som helst stoppa transkriberingen (om t.ex. mötet avslutas tidigare) genom att åter klicka på Fler åtgärder > Spela in och transkribera > Stoppa transkription. Om du glömmer att stoppa manuellt avslutas transkriberingen automatiskt när alla lämnat mötet.
7. **Efter mötet – hitta transkriptionen:** Så fort mötet är slut sparas den fullständiga transkriptionen i mötets chatt/logg. Du kan hitta den genom att öppna Chat-fliken i Teams och klicka på den avslutade mötets chattrum. I chatten syns ett meddelande om att "Mötet avslutades" med en länk till Transkription – klicka där för att öppna texten. (Alternativt kan du gå via kalendern: öppna mötesevenemanget i Teams-kalendern, där transkriberingen finns under Sammanfattning.)
8. **Ladda ner transkriptionen:** För att spara en kopia, gå till fliken Sammanfattning (Recap) uppe i möteschatten eller mötesdetaljerna. Klicka på nedladdningspilen (↓) vid transkriptionen och välj format – antingen .docx (Word) eller .vtt (VTT subtitle) fil. Spara filen på din dator. (Endast mötesorganisatören och medorganisatörer kan ladda ner transkript som standard, men admin kan ändra behörigheter om fler behöver åtkomst.)

Förinställd transkribering: Som organisatör kan du i Mötesalternativ (innan mötet startar) aktivera "Spela in och transkribera automatiskt". Detta gör att Teams per automatik börjar spela in och transkribera direkt när mötet startar – praktiskt om du vet att du annars kan glömma att slå på det manuellt. Kontrollera dock att alla deltagare är medvetna om att mötet kommer spelas in.

Passande användningsområden och begränsningar

När fungerar Teams-transkribering som bäst? Den lämpar sig väl för vanliga affärsmöten, presentationer, utbildningar och liknande där en person talar åt gången och ljudet är klart. I sådana scenarier kan transkriberingen ge ett detaljerat, sökbart protokoll över diskussionen utan att någon behöver föra manuella anteckningar. Det underlättar även för frånvarande att i efterhand ta del av mötet, eller för deltagare att

fokusera på samtalet istället för att anteckna. Transkribering ökar dessutom tillgängligheten – t.ex. för personer med nedsatt hörsel – genom att de kan läsa allt som sägs i realtid.

Begränsningar: I mer dynamiska möten – exempelvis brainstorms med många som talar i mun eller diskussioner med bakgrundsljud – kan Teams ha svårare att ge en felfri avskrift. Om flera pratar samtidigt riskerar en del tal att utelämnas eller bli felaktigt transkriberat (ASR-teknologin försöker tolka en röst åt gången). Även specialiserat språk (branschspecifika termer, namn eller förkortningar) kan transkriberas fel och behöva korrigeras i efterhand. För möten med blandade språk kan funktionen bara använda ett språk åt gången – om deltagare skiftar mellan t.ex. svenska och engelska mitt i mötet kan det kräva att man ändrar transkriptionsspråk manuellt för bästa resultat.

Jämförelse med andra verktyg: Jämfört med dedikerade transkriptionslösningar kan Teams erbjuda smidighet snarare än absolut toppprecision. Exempelvis OpenAI Whisper (öppen källkod) eller tjänster som Klang.ai kan uppnå mycket hög noggrannhet och stödjer många språk, inklusive svenska, oftast genom att transkribera inspelade ljudfiler. Whisper är känt för att fånga även detaljer och nyanser väl, men kräver teknisk hantering utanför Teams (t.ex. ladda ned mötesinspelningen och köra en transkribering separat). Microsoft Word har en inbyggd transkriberingsfunktion i webbversionen där man kan ladda upp mötets inspelning; den ger också en text med talarsegment, men är begränsad i antal minuter per månad och måste köras manuellt efter mötet. Fördelen med Teams inbyggda transkribering är integrationen och realtiden – texten genereras direkt under mötet och kopplas till deltagarnas identiteter, något som externa verktyg inte kan göra automatiskt.

Sammanfattningsvis kan man säga att Teams-transkribering är bättre på helheten – det vill säga att snabbt ge en översikt och kontext av mötet med vem som sa vad – medan specialiserade verktyg kan vara bättre om man behöver full detaljprecision (t.ex. ordagranna utskrifter för juridiska protokoll eller högkvalitativa intervjuer). Teams strävar efter att få med det mesta korrekt, men vissa detaljer kan behöva verifieras; alternativt som Whisper kan träffa rätt på fler ord men saknar Teams smidiga samarbetskontext (och kräver mer efterarbete för att identifiera talare). Med Teams Premium-tillägget får man dessutom AI-genererade sammanfattningar och action items efter mötet, men själva transkriptionsmotorn är densamma som i vanliga Teams. Det innebär att Premium ger mervärde i form av smarta insikter (helheten) snarare än bättre ordagrann noggrannhet.

Identifiering av röster och efterarbete

En stor fördel med Teams inbyggda transkribering är att den identifierar olika talare automatiskt. Varje gång en person pratar i mötet lägger transkriberingen in deras namn före utskriften av deras tal (t.ex. "Anna: [tal]"). Detta bygger på mötesdeltagarnas Teams-identiteter – så länge personen är med i mötet (och inte har valt att anonymisera sig)

kommer deras namn visas. Om någon däremot ansluter som gäst eller via telefon utan identifikation kan transkriberingen visa mer generiska etiketter (eller ingen namnlabel alls). Notera att deltagare har möjligheten att dölja sin identitet i transkriptionen om de önskar, av integritetsskäl, men då står det exempelvis "Deltagare 1" istället för namnet. På det stora hela ger Teams transkribering en tydlig uppdelning per talare, vilket är värdefullt för att följa dialogen i efterhand.

Efterarbete: Även om transkriptionen sker automatiskt behöver man räkna med visst efterarbete för att få en helt korrekt och läsbar slutprodukt. För det första är det klokt att gå igenom texten efter mötet och rätta eventuella felaktigheter – t.ex. namn som stavats fel, facktermer som blivit konstiga, eller om något inte transkriberats alls på grund av brus. Teams gör automatiskt grundläggande formatering (t.ex. meningsbrytningar och punkter), men ibland kan meningsindelningen behöva justeras för läsbarhet. Du kan öppna transkriptionen i Word för enkel redigering om du valt .docx-formatet. För det andra kan det vara bra att sammanfatta eller extrahera beslut ur transkriptet manuellt. En ordagrann avskrift kan bli lång; ofta vill man dra ut de viktiga punkterna. (Här kan förstås AI-sammanfattning i Teams Premium eller Copilot vara till hjälp om tillgängligt, men annars får man göra det manuellt.) Ibland kan även tidsstämplarna behöva tas bort om man vill presentera texten som ett rent protokoll – de finns där för referens men kanske inte behövs i en slutrapport. Slutligen, se över att talartilldelningen stämmer hela mötet igenom; om två personer pratade nästan samtidigt kan det hända att något stycke hamnat under fel namn, vilket man då kan korrigera. Genom att lägga några minuter på dessa justeringar kan man gå från en rå transkription till en tydlig och pålitlig dokumentation av mötet.

Sammanfattande rekommendation

Microsoft Teams transkribering passar bäst för organisationer som redan arbetar i Teams och vill ha ett enkelt, integrerat sätt att spela in mötesdialogen skriftligt. Den är idealisk för att förbättra möteseffektivitet och inklusivitet – alla deltagare kan följa med och efteråt finns en sökbar logg. Om ni värdesätter att ha talaridentitet och realtidsåtkomst till transkripten direkt i samarbetsplattformen, är Teams-transkribering ett utmärkt val. För de flesta vanliga möten ger den tillräckligt bra resultat, särskilt på engelska och större språk, och sparar mycket tid jämfört med manuell transkribering. Kort sagt bör ni välja Teams inbyggda transkription om ni vill ha snabb och smidig mötesdokumentation inom Microsoft 365-miljön, utan att behöva ta till externa verktyg.

Vanliga nybörjarmisstag och tips för att undvika dem

- **Glömmer att aktivera funktionen:** Ett vanligt misstag är att anta att transkribering sker automatiskt. I själva verket måste man starta den manuellt under mötet (om man inte ställt in automatisk start). Tips: Sätt en påminnelse i agendan att klicka på "Starta transkription" i början, eller använd mötesalternativet för automatisk transkribering så slipper du komma ihåg det.

- **Saknar rätt behörighet/licens:** Nya användare kan bli förvånade om transkriberingsknappen är grå. Ofta beror det på att man är inloggad med fel konto eller att admins inte aktiverat funktionen. Till exempel kan gäster och externa deltagare inte starta transkribering – bara någon i organisatörens organisation kan göra det. Även anonyma deltagare (oinloggade) saknar tillgång. Tips: Se till att du är inloggad med ditt arbetskonto och att du antingen är mötesorganisatör eller presentatör. Om knappen säger att funktionen inte är tillåten enligt organisationsprincip, kontakta IT-administratören – det kan krävas att de slår på transkribering i policyinställningarna för möten.
- **Försöker transkribera på fel typ av möte:** Som nämnt fungerar det endast i schemalagda möten. Om du startar ett snabbmöte (Meet now) utanför kalendern eller ett rent Teams-samtal (utan mötesinbjudan) kommer transkriberingsalternativet inte finnas. Tips: Boka in ett möte i förväg (det går bra att göra det direkt före om det är ad-hoc, men använd kalendern) så att alla Teams-funktioner, inkl. transkribering, är tillgängliga.
- **Fel språk inställt:** Om transkriberingen verkar väldigt inkonsekvent kan det bero på att fel språk är valt (t.ex. engelska som standard, medan mötet hålls på svenska). Tips: Kontrollera språkvalet när du startar transkriberingen – Teams brukar försöka upptäcka språk men om det inte stämmer, ändra till rätt språk för bättre resultat.
- **Orealistiska förväntningar på noggrannhet:** Nybörjare kan anta att 100% av orden blir perfekt avskrivna. I praktiken är automatiska transkriberingar mycket hjälpsamma men inte felfria. Tips: Använd transkriptionen som ett stöd och utgångspunkt, men korrekturläs alltid om innehållet är affärskritiskt. För viktiga möten kan man överväga att kombinera Teams-transkriptet med en snabb manuell översyn eller använda en professionell tjänst om ordagrann exakthet krävs.

Genom att vara medveten om dessa vanliga fallgropar och följa tipsen ovan kan ni undvika frustrationen och få ut maximal nytta av transkriberingsfunktionen i Microsoft Teams. Med lite förberedelse och efterkontroll blir det ett kraftfullt verktyg för att dokumentera och dela kunskap från era möten.

Hur använda KB Whisper – lite mer avancerat men så väldigt bra!

Kort introduktion

KB-Whisper är en AI-modell för tal-till-text-transkribering utvecklad av Kungliga bibliotekets KBLab. Modellen baseras på OpenAI:s Whisper men är extra tränad på över 50 000 timmars svenskt tal, vilket ger hög noggrannhet när den omvandlar tal (på svenska eller engelska) till skriven text.

Förutsättningar och fakta

- **Licens och kostnad:** Modellen är helt gratis att ladda ner och använda (öppen källkod under Apache 2.0-licens). Inget abonnemang eller betaltjänst krävs.
- **Enheter:** KB-Whisper körs främst lokalt på dator (Windows, Mac eller Linux) via en Python-miljö. Det finns inget grafiskt användargränssnitt eller mobilapp i nuläget – man interagerar med modellen genom kod eller kommandon. (Det går att bygga in modellen i en egen webb- eller mobilapplikation, men det finns ingen färdig officiell app.)
- **Språk och noggrannhet:** Whisper-teknologin stöder många språk (grundmodellen kan transkribera ~97 språk med hög kvalitet). KB-Whisper är särskilt optimerad för svenska, men kan även transkribera engelska och andra språk. För svenska når modellen mycket hög noggrannhet – enligt tester ca 47% färre fel på ordnivå jämfört med original-Whisper från OpenAI. Den hanterar svenska dialekter och uttal väl tack vare det omfattande träningsdatat och är exempelvis duktig på att stava svenska namn och platser korrekt.

Steg-för-steg-instruktion för transkribering

Följ dessa steg för att installera och använda KB-Whisper lokalt på din dator. Inga förkunskaper om AI krävs, och vi går metodiskt igenom varje moment:

1. **Installera Python och öppna terminalen:** Se till att du har Python 3 installerat på datorn (kan laddas ner från python.org om det saknas). Öppna sedan en terminal/kommandotolk (t.ex. PowerShell i Windows eller Terminal i macOS/Linux) där du kan köra kommandon. Detta är nödvändigt för att installera paket och köra transkriberingsskriptet.
2. **Installera nödvändiga Python-paket:** I terminalen, installera Hugging Face Transformers och PyTorch (eller annan kompatibel AI-ramverk) samt eventuellt ljudhanteringsbibliotek. Du kan göra detta med pip. Till exempel:
3. `pip install transformers datasets torch soundfile`

Paketet transformers innehåller verktyg för att ladda modellen, torch är PyTorch (krävs för att köra modellen), och soundfile hjälper till att läsa ljudfiler. Har du en NVIDIA-GPU och vill utnyttja den, se till att installera en version av PyTorch med CUDA-stöd (följ instruktionerna på pytorch.org). Tips: Om pip inte fungerar eller du får fel, kontrollera att Python är korrekt installerad och att du kör kommandot i rätt terminal.

4. **Ladda ner KB-Whisper-modellen:** Du behöver nu ladda modellen från Hugging Face. Detta kan ske automatiskt via skriptet – du behöver bara ange modellens namn så sköts nedladdningen åt dig första gången. I Python kan du till exempel skriva:

5. `from transformers import pipeline`
6. `# Välj modellstorlek: "kb-whisper-large" för bästa noggrannhet,`
7. `# eller t.ex. "kb-whisper-small" för snabbare transkribering på svagare dator.`
8. `model_name = "KBLab/kb-whisper-large"`
9. `transcriber = pipeline("automatic-speech-recognition", model=model_name)`

Första gången du kör detta laddas modellens viktfiler från Hugging Face (vilket kan vara flera gigabyte för large-versionen – ha tålamod och internetuppkoppling). Du kan byta ut kb-whisper-large mot kb-whisper-small, medium osv. om du vill använda en mindre modell. Om du använder GPU kan du initiera pipelinen med `device=0` (för att använda första GPU:n); annars används CPU som standard (vilket kan vara betydligt långsammare för stora modeller).

10. **Förbered ljudfilen:** Se till att du har din ljudinspelning i ett format som modellen klarar. Vanliga format som WAV, MP3 eller MP4 stöds. Om du använder ett komprimerat format som MP3/MP4 kan det krävas att FFmpeg är installerat i systemet (för att pipeline:en ska kunna läsa filen). Ett säkert alternativ är att använda en WAV-fil (16 kHz, mono) om möjligt, då undviker du formatproblem. Döp filen till något enkelt, t.ex. inspelning.wav, och lägg den i en sökbar mapp (eller notera sökvägen).

11. **Kör transkriberingen:** Nu är det dags att låta modellen transkribera ljudet. Anropa pipeline-objektet med sökvägen till din ljudfil. För svenska transkriberingar är det bra att ange språket uttryckligen så att modellen inte försöker identifiera fel språk. Exempel:

12. `result = transcriber("inspelning.wav", generate_kwargs={"task": "transcribe", "language": "sv"})`
13. `text = result["text"]`
14. `print(text)`

Modellens transkriberingsfunktion läser in ljudfilen och returnerar resultatet som text. I exemplet ovan sparas transkriberad text i variabeln text. (Parametern `language: "sv"` anger att det rör sig om svenska. Om du transkriberar engelskt tal kan du sätta `"language": "en"` eller låta modellen autodetektera.) Observera att KB-Whisper transkriberar inspelat material, alltså färdiga ljudfiler. Transkriberingen sker inte i realtid medan du pratar, utan det är en efterskottsprocess – du kör modellen på en redan inspelad fil. (Realtidsanvändning är möjligt att bygga med kod, men ligger utanför grundläggande användning.)

15. **Spara eller använd resultatet:** När transkriberingen är klar kan du se den utskrivna texten i terminalen. Kopiera texten och spara den i valfri texteditor, eller låt Python spara den direkt till en fil:

16. `with open("transkribering.txt", "w", encoding="utf-8") as f:`

17. `f.write(text)`

Nu har du en textfil `transkribering.txt` med det avlyssnade innehållet. Du kan öppna den i Word eller annat program för vidare redigering. Tips: I mer avancerad användning kan man även få med tidsstämplar för varje textsegment (genom att ange `return_timestamps=True` i pipeline-anropet), men till en början räcker det oftast med ren text.

Passande användningsområden och begränsningar

Användningsområden: KB-Whisper lämpar sig väl för att transkribera inspelade möten, intervjuer, seminarier och workshops – i princip alla situationer där du har ljud med tal som du vill få nedskrivet. Det kan vara projektmöten som spelats in via Zoom/Teams, intervjuer för forskning eller media, podcasts, videoklipp eller föreläsningar. Modellen har också visat sig användbar inom undertextning av TV-sändningar, hantering av kundtjänstsamtal och utskrift av diktat inom sjukvården. Eftersom KB-Whisper kan köras offline på egen dator passar den bra när man behöver transkribera känsligt material utan att ladda upp ljudet till en molntjänst – till exempel interna möten eller intervjuer med konfidentiellt innehåll.

Begränsningar: Det krävs lite teknisk handlag för att använda KB-Whisper (eftersom den saknar klickbart gränssnitt). Till skillnad från t.ex. Microsoft Teams som erbjuder live undertexter/medverkandelistor under ett möte, arbetar KB-Whisper efter mötet på inspelningen. Teams och liknande verktyg kan ge snabba realtidsundertexter, men de kan missa en del detaljer och kräver internetanslutning samt ofta ett företagskonto. Microsoft Word (webbversion via Office 365) har en inbyggd transkriberingstjänst som automatiskt identifierar olika talare och transkriberar ljud du laddar upp – det är användarvänligt men kräver ett abonnemang och har begränsningar i filstorlek/uppladdningstid. Jämfört med dessa kan KB-Whisper erbjuda högre noggrannhet på svenska (särskilt för namn, platser och facktermer) eftersom modellen är specialtränad på svenskt tal. Original-Whisper från OpenAI (och andra molntjänster) kan absolut transkribera svenska ganska bra, men KB-Whisper gör färre fel och hanterar svenska dialekter bättre. När det gäller helhet vs. detaljer i transkriptionen får du med KB-Whisper både en god helhetsåtergivning och fler korrekta detaljer – modellen är mindre benägen att höfta eller gissa fel på specifika svenska ord. En möjlig begränsning är att transkriberingen kan ta lite tid för långa ljudfiler, särskilt på en dator utan kraftfull GPU. Sammanfattningsvis är KB-Whisper bäst när du prioriterar noggrannhet och

kontroll över transkriberingen, medan verktyg som Teams eller Word kan vara smidigare för omedelbarhet och enkelhet (dock på bekostnad av viss precision).

Identifiering av röster och efterarbete

En viktig sak att känna till är att KB-Whisper inte automatiskt skiljer mellan olika talare i en inspelning. Modellen transkriberar allt tal som en sammanhängande textmassa. Det betyder att om flera personer pratar i ljudfilen kommer utskriften inte ange byten mellan talare eller vem som sa vad. För mötesanteckningar kan detta vara en begränsning – du får manuellt gå igenom texten i efterhand och markera talarbyten (t.ex. skriva ut namnet på personen före deras uttalanden) baserat på minnet av mötet eller genom att lyssna på inspelningen igen.

Efterarbete kan alltså behövas för att få ett tydligt slutresultat. Ett vanligt efterarbete är att markera talaridentitet: lägga till etiketter som [Person A]: när en ny röst tar vid, särskilt om dialogen är viktig att attribuera. Du kan underlätta detta genom att notera tidsstämplar; om du använde funktionen för tidsstämplar i modellen kan du se var byten sker och infoga radbrytningar där. Andra efterbearbetningar inkluderar att rätta till eventuella felaktiga ord som modellen missförstod (även om KB-Whisper är noggrann kan bakgrundsljud eller överlappande tal orsaka fel som du behöver justera manuellt). Du kan även behöva justera formatering – modellen sätter ut skiljetecken och stor bokstav i meningar automatiskt, men längre utdrag kan behöva styckindelas för läsbarhet. Sammanfattningsvis får du räkna med att läsa igenom och redigera transkriptionen efteråt för att identifiera vem som sa vad och säkerställa att allt är korrekt och tydligt. För professionellt bruk (t.ex. utskrift av en intervju) är det klokt att kombinera modellens hastighet med en mänsklig genomläsning för bästa resultat.

Sammanfattande rekommendation

Vem bör välja KB-Whisper och varför? Denna modell passar bäst för den som behöver hög precision i svensk taligenkänning och föredrar en lokal, kostnadsfri lösning över molnbaserade tjänster. Om du ofta arbetar med svenska ljudinspelningar – exempelvis journalist som vill transkribera intervjuer, forskare som analyserar inspelade samtal, eller ett företag/myndighet som vill protokollföra möten internt – är KB-Whisper ett utmärkt val då det kan hantera svenska språkets detaljer bättre än de flesta alternativ. Du får full kontroll över data (inget lämnar din dator) och slipper licenskostnader, vilket gör modellen värd att använda för många som är villiga att lägga några minuter på installation och efterarbete.

Vanliga nybörjarmisstag och tips för att undvika dem

- **Glömmer att installera Python eller kör i fel miljö:** Ett vanligt misstag är att försöka köra kommandon utan att ha installerat Python, eller i fel fönster. Lösning: Installera alltid Python 3 först och använd rätt terminal (t.ex. Anaconda

Prompt om du använder Anaconda, eller se till att python/pip kommandon fungerar i ditt skal). Kontrollera genom `python --version` i terminalen.

- **Missar att installera rätt paket:** Om du bara laddar ner koden men inte installerar Hugging Face Transformers eller PyTorch, kommer inget att fungera. Lösning: Följ steg 2 ovan noggrant – kör `pip install ...` kommandot för alla nödvändiga bibliotek. Dubbelkolla stavningen av paketnamn. Om installationen krånglar på Windows, prova att köra kommandot som administratör eller via Anaconda.
- **Använder fel modellnamn eller variant:** KB-Whisper finns i olika storlekar (tiny, base, small, medium, large). Nybörjare kan bli förvirrade och kanske försöker använda OpenAI:s originalmodell av misstag eller fel variant. Tips: Se till att du anger `KBLab/kb-whisper-*` i modellnamnet. Börja gärna med `kb-whisper-small` om din dator saknar kraftfull GPU – det går snabbare och kräver mindre minne. Du kan alltid byta till `kb-whisper-large` senare för högre noggrannhet.
- **Förväntar sig realtidsresultat:** Vissa tror att modellen ska transkribera direkt medan de pratar (som t.ex. Siri eller Teams live-text). Standard-KB-Whisper gör inte detta automatiskt. Förklaring: Modellen transkriberar efter att du har en färdig ljudfil. För live-transkribering måste man programmera en lösning som matar in ljudströmmar fortlöpande, vilket är mer avancerat. Som nybörjare bör du fokusera på inspelade filer.
- **Ingen GPU – allt går långsamt:** Om du kör den största modellen på en vanlig laptop utan GPU kan transkriberingen ta väldigt lång tid eller eventuellt fylla ditt RAM. Tips: använd mindre modeller eller se till att du har åtminstone en basalt kraftfull dator. Alternativt kan du köra transkriberingen i en molnmiljö som Google Colab (som ofta har GPU) om din egen dator är för långsam.
- **Glömmer att spara eller säkerhetskopiera resultatet:** Efter att modellen printat ut texten i terminalen kan det vara frestande att bara stänga fönstret. Då förlorar du transkriptionen. Lösning: Spara alltid ut resultatet, antingen genom att omdirigera output till en fil eller genom att kopiera och klistra in texten i ett dokument. Du kan också använda Python (som visat ovan) för att skriva direkt till en textfil för att undvika misstag.
- **Ljudfilen kunde inte läsas:** Ett annat vanligt problem är att modellen klagar på att den inte kan läsa filen eller att det är fel format. Orsak: Kanske saknas FFmpeg eller ljudbibliotek, eller filvägen är fel. Tips: Använd WAV-format om möjligt. Om du får fel med MP3/MP4, installera ffmpeg eller soundfile (som i steg 2). Kontrollera också att du anger rätt sökväg till filen – om Python-skriptet ligger i en annan mapp, kanske du måste skriva fullständig sökväg (t.ex.

"C:\Users\MittNamn\Dokument\mote.mp3" på Windows med dubbla backslash).

Genom att vara medveten om dessa fallgropar och följa tipsen ovan kan du undvika de vanligaste problemen. Med rätt förberedelser blir användningen av KB-Whisper relativt okomplicerad och väldigt givande – du får snabb, korrekt transkribering av tal på svenska direkt på din egen dator, utan krångel.

Använda Microsoft Words inbyggda transkriberingsfunktion

Kort introduktion

Microsoft Words transkriberingsfunktion låter dig automatiskt omvandla tal (t.ex. från möten eller inspelningar) till text direkt i Word. Funktionen använder molnbaserad AI för att skilja ut olika talare och skapa en tidsstämplad texttranskription av konversationer.

Förutsättningar och fakta

- **Licenskrav:** Du behöver en aktiv Microsoft 365-prenumeration för att använda transkribering i Word. Det är en premiumfunktion i Word för Microsoft 365 (ingår ej i fristående Word utan abonnemang).
- **Internetanslutning:** Funktionen kräver internet eftersom talet bearbetas i molnet. Se till att du är inloggad i Word med ditt Microsoft-konto och har OneDrive tillgängligt (inspelningar sparas automatiskt på OneDrive).
- **Enheter och appar:** I skrivande stund fungerar transkribering endast i Word-appen för Windows (Microsoft 365-versionen) samt via den mobila Microsoft 365-appen (Office-appen) på Android/iOS. Den är inte tillgänglig i den fristående Word-appen på mobil/surfplatta eller i Word för Mac ännu (där får man använda Word Online som alternativ).
- **Stödda språk:** Transkriberingsfunktionen kan hantera över 80 olika språk och språkområden. Både svenska och engelska stöds, liksom många andra språk. Du kan välja språk i transkriberingsmenyn för bästa resultat.
- **Noggrannhet:** För tydligt tal på huvudspråken (t.ex. engelska) kan transkriberingen vara relativt korrekt, men det är sällan felfritt. På svenska fungerar det också, men var beredd på att behöva rätta vissa ord och lägga till skiljetecken manuellt (se mer under Efterarbete). Vid oklar eller dialektal tal kan fel förekomma. Generellt ger verktyget "huvudinnehållet" rätt, men det kan missa detaljer eller tolka vissa ord fel.
- **Tidsbegränsning:** Det finns en gräns på 300 minuter per månad (5 timmar) för automatiska transkriberingar per användare. Denna gräns gäller uppladdat ljud (och inspelningar) och kan inte höjas. Planera användningen så att dina viktigaste möten ryms inom denna kvot.

Steg-för-steg-instruktion för transkribering

Följ dessa steg för att transkribera ett möte eller en ljudinspelning i Word:

1. **Öppna Word och hitta funktionen:** I Word (Microsoft 365) på datorn, öppna ett nytt eller befintligt dokument. Gå till Start-fliken i menyfliksområdet, klicka på droppmenyn vid knappen Diktera (mikrofonikonen) och välj Transkribera. (Första gången kan Word be om tillåtelse att använda mikrofonen och anslutna onlinetjänster – godkänn detta för att fortsätta.)
2. **Välj inspelning eller uppladdning:** I panelen Transkribera som öppnas till höger har du två alternativ: Starta inspelning (för att spela in live tal) eller Ladda upp ljud (för att transkribera en redan inspelad fil).
 - **För att spela in direkt:** Klicka Starta inspelning. Word börjar spela in via din dators mikrofon. (Tips: kontrollera att rätt mikrofonkälla är vald i datorns ljudinställningar, så att ljudet blir tydligt.) När inspelningen pågår ser du en tidsräknare; ingen text visas ännu under inspelningen eftersom transkriberingen sker i bakgrunden. Tala tydligt och genomför mötet/intervjun som vanligt. Du kan pausa inspelningen med mikrofonknappen och återuppta vid behov.
 - **För att transkribera en fil:** Klicka Ladda upp en ljudfil. Välj din inspelade fil via dialogrutan. Word stöder formaten .wav, .mp4, .m4a och .mp3. (Observera att eventuella videofiler måste konverteras till något av dessa ljudformat först – Word transkriberar bara ljudspåret.)
3. **Avsluta och spara:** Om du spelat in live, tryck på Spara och transkribera nu när mötet är klart. Inspelelingen laddas då upp till OneDrive och den automatiska transkriberingen startar. (Om du laddat upp en fil sker transkriberingen direkt efter uppladdning.) Du kan fortsätta arbeta i Word eller annan app under tiden, men låt transkriberingsfönstret vara öppet i bakgrunden. Transkriberingen tar ungefär lika lång tid som ljudets längd (beroende på internet).
4. **Granska resultatet:** När processen är klar visas texten i Transkribera-panelen, uppdelad i avsnitt med tidsstämplar. Varje gång en ny röst talar kommer det stå Talare 1, Talare 2 etc. vid avsnittet. Scrolla igenom transkriptionen medan du spelar upp ljudet om du vill kontrollera – du kan klicka på en viss textpassage och välja uppspelningsknappen för att höra just den delen av ljudet (Word markerar texten som spelas upp).
5. **Redigera och exportera text:** Du kan nu redigera transkriptionen direkt i panelen för att rätta eventuella fel. Om Word har märkt upp talare som Talare 1 etc., kan du klicka på namnet och byta ut det (t.ex. till personens riktiga namn); du kan även välja "Ändra alla" för att ersätta en talaretikett i hela dokumentet. Kontrollera

språket och lägg till skiljetecken där det behövs. När du är nöjd kan du infoga texten i dokumentet: klicka på Lägg till i dokument i Transkribera-panelen. Du får då alternativ att antingen infoga hela transkriptionen som text, eller bara utvalda delar eller höjdpunkter. Du kan också välja att Spara hela transkriberingen som ett separat Word-dokument om du vill behålla den som referens.

Är transkriberingen i realtid? – Både ja och nej. Under själva inspelningen i Word visas inte texten löpande (det är inte som live-textning på skärmen). Transkriberingen genereras först efter att du sparar inspelningen eller laddat upp filen, alltså efter mötet. I mobilappen (Microsoft 365-appen) finns däremot en funktion för "live transcription" som visar text medan du spelar in, men även där sparas slutresultatet för efterhandsbruk. Sammanfattningsvis: Du får den färdiga texten efter att inspelningen avslutats och bearbetats – inte i realtid under mötets gång.

Passande användningsområden och begränsningar

Microsoft Words transkriberingsfunktion är användbar i många sammanhang, men lämpar sig bäst för små till medelstora möten eller intervjuer där talet är hyfsat tydligt. Här är några exempel och överväganden:

- **Intervjuer och samtal:** Perfekt för att spela in en intervju, ett samtal eller en podcast och få dialogen utskriven. Word kan hantera att flera personer pratar och separerar talarna i texten. Du kan även föra egna anteckningar i dokumentet samtidigt som du spelar in, vilket är praktiskt vid t.ex. forskningsintervjuer eller journalistik.
- **Möten och workshops:** För interna projektmöten eller brainstorming-sessioner kan transkribering ge en helhetslogg av vad som sades. Alla deltagare behöver inte fokusera på att ta protokoll, eftersom man i efterhand kan gå igenom transkriptet för att plocka ut beslut och uppgifter. Tänk på att väldigt stora möten med många som pratar i munnen på varann blir svåra att transkribera korrekt – AI:n gör inte alltid ett bra jobb om flera talar samtidigt eller om bakgrundsljud stör.
- **Föreläsningar och presentationer:** Om du spelar in en föreläsning eller webinar kan Word hjälpa dig att få talet i textform. För ensamtal (en talare åt gången) blir kvaliteten oftast bättre än för gruppdiskussioner. Även om transkriberingen kanske missar något ord här och där kan den ge ett utmärkt utgångsmaterial för att skapa sammanfattningar, mötesminnen eller undertexter.

Jämförelse med andra verktyg: Det finns alternativ för automatisk transkribering, och vilket som är bäst beror på behov:

- **Word vs. KB Whisper:** KB-Whisper är en svensk AI-modell baserad på OpenAI Whisper, specialtränad på svenska för hög precision. Den kan ge mer noggranna

transkriberingar på svenska än Word och körs dessutom lokalt (ingen internetanslutning behövs, vilket är bra för känsligt material). Å andra sidan kräver KB Whisper att man använder ett separat program eller script – integrationen är inte lika smidig som Word. För en van användare som behöver maximal precision och sekretess på svenska kan KB Whisper vara bättre, men för enkelhetens skull vinner Word om du redan arbetar i Office-miljön.

- **Word vs. Microsoft Teams:** Microsoft Teams har en inbyggd funktion för live-transkribering av möten (med stöd för vissa språk). I ett Teams-möte kan man slå på transkribering så att allt som sägs visas i realtid som text, kopplat till respektive talares namn (eftersom Teams vet vem som pratar). Efter mötet kan transkriptet sparas i Teams. För möten som ändå hålls i Teams kan detta vara mer praktiskt – du får då talaridentitet automatiskt med namn och behöver inte spela in separat. Nackdelen är att det bara fungerar inom Teams-möten. Word-transkribering är mer allmän – den kan användas för vilket som helst ljud, inte bara Teams-samtal (t.ex. fysiska möten i ett rum, telefonsamtal på högtalare, etc.). Om helhetsdokumentation av ett Teams-möte med flera kända deltagare är målet, kan Teams transkribering vara bättre; om du behöver transkribera fristående ljudklipp eller möten utanför Teams, är Word rätt väg.
- **Word vs. andra tjänster:** Det finns många tredjepartsverktyg (t.ex. Otter.ai, Google Docs röstinmatning, osv). Google Docs erbjuder t.ex. röstskrivning via Chrome som kan transkribera ljud som spelas upp, och det fungerar på svenska också. OpenAI Whisper (öppen källkod) är känt för mycket hög noggrannhet och flerspråkighet, men kräver teknisk kunnighet för att användas fullt ut. En jämförelse lyder: "Om du vill ha bekvämlighet, använd Word eller Google Docs; om du strävar efter högsta precision och bred språkmoderstöd, välj Whisper." Med andra ord – Words styrka är enkelheten och integrationen i Office, medan mer specialiserade verktyg kan ge bättre detaljprecision i vissa fall.

Helheten vs. detaljer: Word är duktig på att fånga helheten i ett samtal – du får snabbt ett fullständigt utkast av allt som sagts, utan att behöva skriva manuellt. Detta är guld värt för att spara tid; även om inte varje ord är perfekt igenget känns essensen av mötet igen. När det gäller detaljer är dock automatiska transkriberingar i Word ibland bristfälliga. Till exempel sätts inte alltid punkter och kommatecken där de borde vara, och en del ord kan bli felaktiga. För ett ordagrant protokoll med hundraprocentig exakthet krävs därför att man går igenom texten och justerar (eller att man använder en professionell transkriberingstjänst manuellt). Så sammanfattat: Word ger dig grunden och helheten väldigt enkelt, men de finstilla detaljerna kan behöva handpåläggning.

Identifiering av röster och efterarbete

En stor fördel med Words transkribering är att den automatiskt försöker särskilja olika talare. När två eller flera personer pratar under ett möte kommer transkriptionen att

delas upp i stycken och varje stycke märks med generiska etiketter som Talare 1, Talare 2, Talare 3, osv. Detta kallas speaker diarization och hjälper dig se när en ny person börjar prata. Notera att Word inte vet vem "Talar 1" egentligen är – du måste själv byta ut etiketterna till namn om du vill (genom att klicka på etiketten i transkriptionspanelen och ändra, eller använda funktionen "Ändra alla" för att byta ut på alla ställen). Om samma person talat hela mötet kommer förstås allt stå under "Talar 1".

Hur bra fungerar talarigenkänningen? I regel identifierar Word byte av talare när det är tydligt paus eller skillnad i röst. Vid två personer runt ett bord fungerar det oftast bra. Om flera personer avbryter varandra eller pratar samtidigt kan AI:n missa och fortsätta samma talar-ID trots att någon annan börjat – dessa får man rätta i efterhand. Du kan också manuellt dela upp eller sammanfoga transkriptionsavsnitt om du märker att den delat fel.

Efterarbete och korrigeringar: Microsoft Word når inte 100% noggrannhet, så räkna med att göra lite manuellt efterarbete för ett riktigt tydligt resultat. Några vanliga moment att fixa efter transkribering:

- **Skiljetecken och styckeindelning:** Ofta kan texten komma ut som långa sjok utan punkt eller kommatecken. Du behöver läsa igenom och sätta punkt där meningar slutar, kommatecken för andningspauser etc. Ibland kan du nästan "höra" i texten var paus bör vara – använd även den inspelade ljudfilen som referens (du kan spela upp direkt i Word-panelen vid de ställen du är osäker på).
- **Korriger felhörda ord:** Ord som uttalas otydligt eller egennamn kan bli fel. Jämför med eventuella anteckningar eller minne från mötet och justera orden. Teknisk terminologi eller ovanliga namn behöver nästan alltid rättas manuellt. I vårt test har t.ex. "Åke" kunnat tolkas som "och" – sådana fel märker man vid genomläsning.
- **Lägg till format eller markeringar:** Transkriberingen ger rå text. Om du behöver ett läsbart protokoll kanske du vill lägga till talarens namn i fetstil före varje stycke, eller formatera texten till stycken/citat. Du kan kopiera ut transkriptionen till huvud dokumentet och där formatera fritt.
- **Kontrollera tidsstämplar:** Word anger tidsstämplar (t.ex. [00:05:30]) vid vissa intervall i transkriptionspanelen. Dessa följer med om du infogar hela transkriptionen. Du kan behålla dem om det är viktigt att veta när under mötet något sades, eller ta bort dem för ett renare textflöde.
- **Översättning (vid behov):** Om mötet var flerspråkigt och du behöver allt på ett språk kan du efter transkribering använda översättningsfunktioner. Notera dock att automatisk översättning av en redan automatisk transkription kan ge kvalitetsförlust – var extra noggrann med kontroll i sådana fall.

I praktiken bör du alltid läsa igenom hela transkriptionen och jämföra med originalinspelningen för att säkerställa att allt är korrekt återgivet. Transkriberingen i Word blir många gånger felaktig just för att programmet inte alltid sätter punkt vid meningar och vissa ord blir fel, men med en genomläsning och redigering går det att få ett mycket bra resultat på kanske halva tiden mot att skriva av allt manuellt.

Sammanfattande rekommendation

Microsoft Words inbyggda transkribering är ett utmärkt verktyg för användare med Microsoft 365 som vill spara tid på att skriva mötesanteckningar. Den rekommenderas för studenter, professionella och andra som behöver transkribera svenska eller engelska möten och intervjuer på ett smidigt sätt. Däremot, om absolut högsta noggrannhet krävs eller om du har väldigt specifika transkriberingsbehov, kan det vara värt att kombinera Word med manuellt efterarbete eller titta på mer specialiserade lösningar. För de flesta vardagliga behov erbjuder dock Word en mycket tidsbesparande helhetslösning direkt i ett verktyg du redan använder.

Vanliga nybörjarmisstag och tips för att undvika dem

- **Misstag:** Att anta att funktionen finns tillgänglig överallt (t.ex. i alla versioner av Word). **Tips:** Kom ihåg att endast Word med Microsoft 365-licens stöder transkribering. Om du öppnar Word-appen på mobilen eller en äldre Word 2019-version och letar efter "Transkribera" kommer du inte hitta den. Använd Word 365 på Windows eller Office-mobilappen som nämnts ovan.
- **Misstag:** Att förväxla Diktera med Transkribera. **Tips:** Word har både dikteringsfunktion (skriv medan du pratar) och transkribering (spela in eller ladda upp och få text efteråt). Se till att du använder rätt knapp. Om du klickar på vanliga mikrofonikonen "Diktera" kommer Word att försöka skriva löpande text som du talar (och den funktionen stöder inte svenska i alla versioner). Transkribering däremot finns under den lilla pilen bredvid mikrofonen där du kan välja "Transkribera". Transkribering visar ingen text förrän du är klar och sparar ljudet, till skillnad från diktering.
- **Misstag:** Att försöka transkribera på mobilen via fel app. **Tips:** Den separata Word-appen på t.ex. iPhone har inte transkriberingsfunktionen inbyggd. Om du vill använda mobilen, ladda ner Microsoft 365 (Office)-appen och använd röstanteckningsfunktionen där. Alternativt, spela in på mobilen med valfri app och ladda upp ljudfilen till Word på datorn i efterhand för transkribering.
- **Misstag:** Att inte ta hänsyn till 300-minutersgränsen. **Tips:** Planera vilka möten som ska transkriberas. Om du t.ex. har fler än 5 timmar möten per månad att transkribera, prioritera de viktigaste. Håll koll på ungefär hur många minuter du använt (Microsoft räknar ihop all din transkriptionstid per månad). Om du når gränsen kommer nya försök att transkribera inte att fungera förrän nästa månad.

Ett alternativ om du tillfälligt passerar gränsen är att använda en annan tjänst (t.ex. gratis Whisper lokalt) för resterande ljud den månaden.

- **Misstag:** Att tro att transkriberingen blir helt perfekt utan redigering. **Tips:** Var beredd på att gå igenom texten. Vanliga korrigeringar är att lägga till punkt/komma, korrigera namn eller facktermer, och ibland dela upp texten i stycken för bättre läsbarhet. En snabb genomgång med ljudet som referens höjer kvaliteten markant. Automatisk transkribering är ett stöd, inte det slutliga dokumentet. Tänk på det som ett utkast som du behöver finputsas.
- **Misstag:** Att glömma säkerhet och integritet. **Tips:** Även om det är smidigt, ladda inte upp känsliga inspelningar hur som helst. Microsoft sparar visserligen inte ljudet permanent på sina servrar efter transkribering, men ljudfilen ligger kvar i din egen OneDrive. Se till att inga obehöriga har åtkomst till din OneDrive-mapp "Transkriberade filer". Radera inspelningen från OneDrive om den innehåller känslig information du inte behöver spara. För extra känsliga samtal kan en offline-lösning (som KB Whisper lokalt) vara att föredra.

Med dessa tips i åtanke kommer du igång med Microsoft Words transkriberingsfunktion på rätt sätt och slipper vanliga fallgropar. Lycka till med dina transkriberingar!

Automatisk transkribering av möten med Klang.ai

Kort introduktion

Klang.ai är ett svenskt AI-verktyg som automatiskt omvandlar tal från möten och intervjuer till text. Det fungerar som en digital mötesassistent som kan ansluta till dina Teams-, Zoom- eller Google Meet-samtal och generera noggranna transkriberingar och sammanfattningar åt dig.

Förutsättningar och fakta

- **Licenser och kostnad:** Du kan börja använda Klang.ai gratis. Det finns en kostnadsfri plan som tillåter obegränsat antal transkriberingar (men med begränsad export) och upp till 10 AI-svar per månad. För avancerade behov finns Professionell-abonnemang (~350 kr/mån efter en 7 dagars testperiod) som ger snabbare transkriberingar, obegränsade exporter, talaridentifiering m.m. Ingen installation krävs – tjänsten körs via kontot online i webbläsaren.
- **Enheter och plattformar:** Klang.ai används i webbläsare på dator (Windows, Mac m.fl.) eller surfplatta. Även via mobilwebbläsare kan du komma åt tjänsten, så du kan exempelvis granska transkriberingar på din telefon. Det finns dessutom en lösning för Android-telefoner som kan spela in och transkribera mobilsamtal automatiskt. Tjänsten integreras smidigt med de vanligaste mötesverktygen – Microsoft Teams, Google Meet och Zoom – och kan agera som en deltagare i mötet.

- **Språkstöd och noggrannhet:** Klang.ai är utvecklad med fokus på svenska och engelska, men stödjer även danska, norska, tyska och många andra språk. Den underliggande AI-modellen är specifikt tränad på svenska språket för högsta möjliga precision. Transkriberingarnas kvalitet uppges vara jämförbar med manuell transkribering av en människa, vilket innebär mycket hög noggrannhet så länge ljudkvaliteten är god. Svenska och engelska tal transkriberas alltså med stor precision, även vad gäller särdrag i språket eller dialekter, och resultatet presenteras i ett redigerbart format.

Steg-för-steg-instruktion för transkribering

Följ dessa steg för att komma igång med automatisk mötestranskribering via Klang.ai:

1. **Skapa konto och logga in:** Gå till Klang.ai och registrera dig (eller logga in om du redan har ett konto). Tjänsten är molnbaserad, så du behöver inte installera något program. Efter inloggning hamnar du i din arbetsyta på Klang.ai.
2. **Koppla din kalender** (frivilligt men rekommenderat): För att låta AI-assistenten automatiskt veta om dina möten, ge Klang.ai åtkomst till din kalender (t.ex. Google Kalender eller Outlook/Office 365 kalender). Detta görs via Klang.ai's inställningar där du godkänner kalenderintegration. När kalendern är kopplad kan tjänsten identifiera möten med videolänkar och förbereda transkribering av dem. (Det går att använda Klang.ai utan kalenderkoppling också, se nästa steg.)
3. **Bjud in Klang.ai till mötet:** När du har ett Teams-/Zoom-/Meet-möte som du vill transkribera, måste Klang.ai's AI-sekreterare bjudas in som deltagare. Detta kan göras på två sätt:
 - **Via kalender:** Lägg till e-postadressen som mötesdeltagare i din kalenderinbjudan. AI-sekreteraren kommer då att få länken och försöker ansluta när mötet startar.
 - **Manuellt via Klang.ai:** Alternativt, öppna Klang.ai och klicka på "Spela in webbmöte" (finns på startsidan eller i ett projekt). I dialogrutan som visas, klistra in mötets URL-länk (Teams/Zoom/Meet-länken) och ge transkriberingen ett namn. Klicka sedan på "Bjud in Klangs AI-sekreterare".

Exempel: Klang.ai's gränssnitt för att spela in ett webbmöte. Här klistrar man in möteslänken (t.ex. Google Meet), anger ett namn och väljer språk för transkriberingen. Därefter klickar man på "Bjud in Klangs AI-sekreterare" för att låta den ansluta till mötet. Assistenten kommer att "knacka på" i mötet och väntar upp till 5 minuter på att bli insläppt.

4. **Starta mötet och släpp in AI-sekreteraren:** När mötet börjar dyker Klang.ai som en virtuell deltagare upp i lobbyn (väntrummet). Du får en notis i mötesverktyget att en ny deltagare vill in – detta är Klangs AI-sekreterare. Godkänn/Släpp in den i

mötet precis som om det vore en vanlig person. (Om du inte släpper in den inom cirka 5 minuter ger den upp försöket att ansluta.)

5. **Automatisk transkribering under mötet:** När AI-sekreteraren väl är med i mötet sköter den inspelning och transkribering helt automatiskt i bakgrunden. Du behöver alltså inte manuellt starta någon inspelning – det räcker att den är med. Under mötets gång kan du fokusera på samtalet; Klang.ai lyssnar på allt som sägs. (Notera att deltagarna inte ser någon löpande text i realtid – transkriberingen blir tillgänglig först i efterhand via Klang.ai)
6. **Hämta transkriberingen:** Så snart mötet avslutats kommer AI-sekreteraren att lämna samtalet. Klang.ai börjar då bearbeta ljudinspelningen. Inom några minuter efter mötet finns transkriberingen redo i ditt Klang-konto. Du hittar den oftast under det projekt eller den lista du valt (t.ex. under "Mina transkriberingar"). Klicka på transkriberingen för att öppna den – där ser du allt tal nedskrivet, ofta med automatisk uppdelning per talare och en sammanfattning eller AI-genererade nyckelinsikter. Du kan nu läsa, redigera och exportera texten.

Transkribering av inspelad ljudfil: Om du istället redan har en inspelning (t.ex. en mp3-fil från ett möte eller en intervju) kan du ladda upp den direkt till Klang.ai. Gå till din arbetsytta och klicka på "Ny transkribering" och välj alternativet att ladda upp en fil. Därefter drar och släpper du ljud- eller videofilen eller väljer den från datorn, och anger vilket språk talet är på. Klang.ai kommer automatiskt att analysera och transkribera filen inom några minuter. Du får sedan ett resultat i samma redigerbara format med talarindelning, precis som vid ett live-möte.

Passande användningsområden och begränsningar

Klang.ai lämpar sig för många situationer där man vill slippa anteckna manuellt:

- **Intervjuer:** Perfekt för jobbintervjuer, forsknings- eller journalistiska intervjuer. AI-sekreteraren kan vara med i onlinemötet och fånga allt som sägs, vilket ger en ordagrann utskrift att analysera i efterhand. Även fysiska intervjuer kan transkriberas genom att man spelar in ljudet och laddar upp det i tjänsten.
- **Möten och workshops:** I projektmöten, workshops eller utbildningar kan Klang.ai agera protokollförare. All dialog sparas automatiskt, så teamet kan fokusera på diskussionen istället för att skriva minnesanteckningar. Efter mötet genereras mötesanteckningar och uppföljningspunkter automatiskt, vilket är till hjälp för att snabbt dela vad som beslöts.
- **Andra användningsområden:** Tjänsten kan användas som diktafon vid fysiska samtal – t.ex. kan du spela in en kundintervju med mobilen och sedan få den transkriberad. Juridiska förhör, styrelsemöten eller föreläsningar kan också

transkriberas för dokumentationens skull. Kort sagt, alla konversationer där innehållet är viktigt att bevara i textform kan dra nytta av Klang.ai.

Jämförelse med andra verktyg: Jämfört med enklare transkriptionsfunktioner som finns i t.ex. Microsoft Word har Klang.ai både bredare språkstöd och fler funktioner. (Microsoft Words inbyggda transkribering stöder främst engelska och har begränsning på ca 300 minuter per månad, medan Klang.ai är optimerad för svenska och kan användas utan tidsgräns i sin gratisversion.) Jämfört med öppna källkods-modeller som KB Whisper – en svensk tal-till-text-modell med hög noggrannhet – erbjuder Klang.ai en mycket mer användarvänlig helhetslösning. Medan KB Whisper kräver teknisk hantering (och bara ger dig rå transkriberad text), så tar Klang.ai hand om hela processen: från att delta i mötet och spela in, till att leverera text, automatiska sammanfattningar och möjlighet att söka/analysera innehållet. Dessutom sker allt säkert via svenska servrar (GDPR-säkrat) i Klang.ai's moln, vilket sparar dig besväret att köra tung AI-lokalt.

Helhet vs. detaljer: Klang.ai fångar både helheten och detaljerna i ett samtal. Transkriptionen är i princip ordagrann (eventuella mindre fel går att rätta i efterhand) och tack vare AI-funktioner får du även automatiskt en sammanfattning och en lista med frågor/svar från mötet. Detta betyder att du snabbt kan greppa huvuddragen i en lång diskussion, samtidigt som du vid behov kan djupdyka i de exakta formuleringarna som användes. Få andra verktyg kombinerar denna noggrannhet med övergripande insikter – många antingen fokuserar på rå text (detaljer) eller bara mötessummeringar. Klang.ai:s styrka är att den levererar båda delar: en precis transkribering samt intelligenta utdrag ur innehållet.

Identifiering av röster och efterarbete

En stor fördel med Klang.ai är att den automatiskt kan särskilja olika talare i transkriberingen. AI-sekreteraren identifierar vem som säger vad under mötet och sätter ut talarnamn vid varje uttalande. Om du bjöd in mötet via kalendern med namngivna deltagare kan transkriberingen direkt visa t.ex. "Anna:" och "Björn:" framför deras repliker, vilket sparar mycket tid jämfört med att manuellt lista ut talare i texten. Denna funktion för talarigenkänning är relativt unik och underlättar läsningen av protokoll enormt. (I gratisplanen kan det finnas vissa begränsningar på talaridentifieringens omfattning, men i betalplanen är det obegränsat.)

Trots den höga automatiseringsgraden kan visst efterarbete vara klokt. Klang.ai's transkribering är mycket träffsäker, men det rekommenderas att du korrekturläser texten snabbt efteråt för att fånga eventuella fel eller otydligheter. Tjänsten hjälper dig här: den markerar potentiella osäkerheter och ger förslag på korrigeringar direkt i editorn. Du kan spela upp ljudet och följa med i texten med en markör, så du lätt hittar rätt ställe att justera. Ofta kan det handla om namn på personer, facktermer eller förkortningar som du kanske vill standardisera i texten. Efter genomgång kan du också komplettera med egna anteckningar eller markeringar om något saknas (även om AI:n genererar en

sammanfattning automatiskt, kanske du vill finjustera den eller lägga till kontext). När du är nöjd kan du exportera transkriberingen till olika format (t.ex. Word, PDF, undertextfiler) eller dela den via en länk. Sammanfattningsvis krävs minimalt efterarbete – mestadels kontroll läsning – tack vare att Klang.ai sköter grovjobbet och till och med attributerar talare åt dig.

Sammanfattande rekommendation

Klang.ai rekommenderas för användare som vill spara tid och få exakta transkriptioner av sina möten eller intervjuer. Särskilt för dig som arbetar på svenska eller flerspråkigt är detta verktyg ovärderligt, eftersom det kombinerar hög precision med smidig integration i dina vanliga mötesrutiner. Många professionella grupper – från forskare och journalister till jurister och HR-personal – kan dra nytta av Klang.ai för att fokusera på samtalets innehåll istället för på att anteckna. Genom att leverera säkra, automatiska mötesanteckningar hjälper Klang.ai dig att inte missa några detaljer och att enkelt dela insikter med teamet.

Vanliga nybörjarmisstag och tips för att undvika dem

- **Inte ge kalenderåtkomst:** Ett vanligt misstag är att hoppa över kalenderintegrationen. Utan den vet inte AI:n om dina möten i förväg. Tips: Koppla din kalender direkt när du börjar använda Klang.ai (du kan alltid begränsa vilken kalender eller vilka möten den ska ha tillgång till). Då kan mötesassistenten planeras in automatiskt. Alternativt, kom ihåg att manuellt bjuda in klang@klang.ai till de möten du vill transkribera.
- **Glömmer att släppa in botten i mötet:** I övrigt att starta mötet kan man missa att klicka "Admit/Släpp in" för Klang.ai-deltagaren som väntar i lobbyn. Om du inte släpper in den kommer ingen inspelning ske. Tips: Håll utkik efter gäst-deltagaren i början av mötet (den heter något i stil med "Klang AI Secretary"). Informera gärna mötets andra deltagare att en AI-assistent ansluter för transkribering, så att ingen av misstag nekar den åtkomst.
- **Fel inställningar för språk:** I vissa fall kan nybörjare ladda upp en ljudfil utan att justera språk, vilket kan leda till sämre resultat om språket identifieras fel. Tips: Se till att språket är ställt till svenska om inspelningen är på svenska (respektive engelska för engelska möten) i gränssnittet när du startar en transkribering. Då används rätt språkmodell för bästa noggrannhet.
- **Osäkerhet kring installation/behörigheter:** Vissa tror att en separat app måste installeras eller att admin-rättigheter krävs i t.ex. Teams. Tips: Klang.ai kräver ingen installation – det är en molntjänst. Om din IT-miljö är låst för externa appar behöver du inte installera något; bjud bara in Klang via e-post till mötet så fungerar det. (För Outlook/Teams kalenderintegration kan organisationens

administratör behöva godkänna appens åtkomst, så kontakta IT om kalenderkoppling skulle strula.)

- **Förlitar sig blint på transkriberingen utan kontrolläsning:** Nybörjare kan anta att 100% korrekthet alltid uppnås. Tips: Även om Klang.ai är väldigt noggrann är det bra att skumma igenom texten efteråt. Korrigera namn och specialtermer vid behov – det går snabbt med de inbyggda verktygen. Då får du ett felfritt dokument att använda eller dela vidare.

Genom att hålla koll på ovan punkter undviker du de vanligaste fallgroparna och kan dra full nytta av Klang.ai för att transkribera dina möten och samtal på både svenska och engelska med minsta möjliga ansträngning.

Tio molnbaserade transkriberingstjänster som stöder svenska

Nedan följer en översikt över de tio största molnbaserade tjänsterna för automatisk transkribering av tal till text som har stöd för svenska. För varje tjänst anges om den också stöder engelska, samt två meningar om dess funktioner, styrkor eller begränsningar.

1. **Google Cloud Speech-to-Text** – stöder även engelska. Tjänsten känner igen över 120 språk och dialekter, inklusive svenska, och kan användas för att transkribera både förinspelat ljud och strömmande tal i realtid. Den använder Googles avancerade neurala nätverksmodeller för hög noggrannhet och har stöd för bl.a. automatisk filtrering av olämpligt språk, men kräver teknisk integration via Googles molnplattform (API eller utvecklarverktyg) för att användas.
2. **Microsoft Azure Speech to Text** – stöder även engelska. Azure erbjuder robust taligenkänning i molnet med stöd för många språk (svenska inkluderat) och möjlighet att anpassa modeller för specifika domäner. Den har realtids-transkribering och andra avancerade funktioner, men i nuläget saknas t.ex. stöd för automatisk talarigenkänning i svenska inspelningar, vilket kan vara en begränsning vid möten med flera deltagare.
3. **Amazon Transcribe** – stöder även engelska. Amazons transkriberingstjänst är en del av AWS och hanterar många språk, däribland svenska, i både realtid och batchläge (uppladdade filer). Tjänsten har funktioner som automatisk talaridentifiering och möjlighet att använda egna ordlistor för bättre träffsäkerhet, och används flitigt i företag för allt från att transkribera kundcentersamtal till att göra mötesinspelningar sökbara.
4. **IBM Watson Speech to Text** – stöder även engelska. IBMs molnbaserade tal-till-text tjänst använder djupinlärningsmodeller och har global språktäckning för flera av de stora språken (t.ex. engelska, arabiska, italienska, kinesiska – svenska stöds också). En styrka är dess möjligheter till anpassning, t.ex. att träna egna

språkmodeller eller använda nyckelordsdetektering, samt hög datasäkerhet för känslig information. Samtidigt är antalet tillgängliga språk något mer begränsat än hos vissa konkurrenter och transkriberingsnoggrannheten kan vara inkonsekvent – verktyget presterar bäst på större språk och kan ha svårigheter att skilja talare åt i vissa fall.

5. **Deepgram** – stöder även engelska. Deepgram är en utvecklarfokuserad API-tjänst för transkribering som erbjuder hög prestanda och stöd för en rad språk, bl.a. svenska, och har successivt utökat språkutbudet (nya modeller stöder 30+ språk inklusive svenska och thai). Tjänsten är känd för sina exakta AI-modeller och snabba transkribering (den används till och med av NASA), men den kräver teknisk integrering (kodning) i egna applikationer och transkriptens kvalitet kan variera mellan olika språk och terminologier – viss manuell efterarbete kan därför behövas för bästa resultat.
6. **Happy Scribe** – stöder även engelska. Happy Scribe erbjuder automatisk transkribering och undertextning i över 45 språk, inklusive svenska, och har även alternativ för manuell (mänsklig) transkribering vid behov. Tjänsten är lättanvänd och uppskattas för hög noggrannhet jämfört med många konkurrenter samt snabb bearbetning – den lämpar sig väl för att transkribera t.ex. möten, intervjuer och video – men den kan ibland ha svårt att perfekt särskilja olika talare i en inspelning.
7. **Sonix** – stöder även engelska. Sonix är en onlineplattform som automatiskt transkriberar ljud och video och den stödjer svenska (tillsammans med 30+ andra språk). Verktyget har funktioner som automatisk talarseparation, autopunktering och sökbar text, samt ger flera exportformat för transkriptioner. Sonix är känt för snabb processing och ett användarvänligt webbgränssnitt för att spela upp och redigera transkript, men saknar till exempel en mobilapp och kan inte göra direktranskribering av live-ljud.
8. **Rev (AI-transkribering)** – stöder även engelska. Rev är en välkänd transkriberingstjänst som numera också erbjuder automatisk AI-transkribering på många språk, däribland svenska (utöver sin klassiska manuella transkriberingstjänst). Dess automattjänst ger snabba resultat (ofta på några minuter) och inkluderar ett interaktivt online-verktyg för att granska och redigera texten samt möjlighet att använda en egen ordlista för facktermer. Rev är uppskattat för hög noggrannhet (den manuella tjänsten garanterar 99% noggrannhet) men som molntjänst har den ingen gratisnivå och erbjuder inte realtids-transkribering – transkriptionerna fås efter uppladdning snarare än omedelbart medan man talar.
9. **Trint** – stöder även engelska. Trint är en AI-driven transkriptionsplattform som hanterar 40+ språk, inklusive svenska. Den är utformad för samarbete och

redigering: användare kan transkribera ljud/video och sedan direkt redigera texten, markera viktiga delar, lägga till kommentarer och exportera i olika format (inklusive undertexter och även översätta transkript till över 50 språk). Trint är kraftfull och snabb, framför allt för engelska (och andra stora språk) med hög transkriptionsnoggrannhet, men precisionen kan vara lägre för mindre språk och verket har viss svårighet att skilja mellan olika talare i en flerpersonskonversation.

10. **Amberscript** – stöder även engelska. Amberscript är en europeisk molntjänst som stöder svenska och över 90 andra språk. Tjänsten erbjuder både automatisk transkribering med AI och möjligheten att beställa manuellt korrigerade transkriberingar för högsta noggrannhet. Amberscript är webbaserad med ett enkelt gränssnitt där användaren kan ladda upp ljud/video och sedan redigera resultatet i en interaktiv texteditor online. En fördel för nya användare är att de första minuterna transkribering är gratis att prova på, varefter prissättningen är pay-as-you-go (per minut) eller abonnemangsbaserad.