

Slutrapport projektgenomförande - Wikimedia Sverige

Projektet är genomfört inom ramen för Post- och telestyrelsens (PTS) innovationstävling Innovation för alla.

Innehållsförteckning

Innehållsförteckning	2
Projektsammanfattning	3
Målgrupp och behov	4
Redogörelse av projektresultatet	4
Beskriv kortfattat projektets mål och leveranser	4
Viktiga insikter	8
Redogörelse av projektets genomförande	9
Genomförandeplan/tidplan	9
Projektets budget	12
Projektets arbetssätt	15
Vad händer nu?	16
Bilaga 1. Målgruppens behov och hinder	19

1 Projektsammanfattning

Organisation/företag:

Wikimedia Sverige

Projekttitel:

Talresursinsamlaren – För ett tillgängligare Wikipedia genom Wikispeech

Projektets start- och slutdatum:

2019-09-01 – 2021-04-30

Kort beskrivning av projektet och dess syfte (max 200 ord):

Talresursinsamlaren är ett initiativ av Wikimedia Sverige, KTH och STTS AB för att bygga en värdefull resurs för talteknologi på svenska och för att göra Wikipedia mer tillgängligt för de som har svårigheter att tillgodogöra sig skriftlig information. Syftet är att vidareutveckla Wikispeech för att kunna lansera Wikispeech talsyntes (hädanefter omnämnt som Talsyntesen) på Wikipedia. En uppsättning verktyg utvecklas för att med hjälp av bl.a. Wikipedias volontärer samla in stora mängder fritt licensierade inspelningar på svenska samt tillhörande annoteringar för att göra Talsyntesen bättre över tid.

Resurserna fyller ett tomrum runt svenska talresurser och kompletterar existerande initiativ. Materialet behövs för utveckling av talbaserade tjänster helt baserade på open source (vilket är ett krav för implementering på Wikipedia). För att uppnå detta arbetar vi för full aktivering av Wikispeech samt förbättrar och utvecklar ny funktionalitet och metodik för crowdsourcing och spridning av fritt licensierade talresurser.

Detta är tänkt som en service till alla som behöver talresurser för sina tjänster. Projektet bidrar därmed till det större ekosystemet av tillgänglighetslösningar.

Kort beskrivning av projektets målgrupp och deras behov (max 200 ord):

Ett stort antal målgrupper har identifierats som på olika sätt kan dra nytta av projektets resultat vid olika tidpunkter. Den första målgruppen är *Wikimediaanvändarna* som kunnat dra nytta av Talsyntesens utveckling då de kan lyssna på den information de är intresserade av. Projektet har också fokuserat mycket på att *forskare* ska ha kännedom och kunna dra nytta av arbetet. Vi har även stött *språkbevarare*, genom att arbeta med minoritetsspråket baskiska samt presenterat projektet på en konferens om minoritetsspråk. Sveriges alla *myndigheter* har kontaktats om projektet. Uppföljande evenemang med företrädare planeras.

I framtiden ser vi också att *återanvändare av verktyget*, *språkstudenter*, *tillgänglighetstjänster*, *språkteknologiutvecklare*, *utvecklare med behov av tal (men ej talteknologer)* samt *enskilda individer* kommer att ha nytta av materialet som tas fram, det är dock en tid bort då aktiv insamling ännu inte

skett. Mjukvaran kan de dock redan nu dra nytta av för egna projekt då den är öppen programvara och vi har lagt mycket resurser på att dokumentera den väl.

2 Målgrupp och behov

Projektet var huvudsakligen inriktat på att nå Wikimediaanvändarna, vilket det varit framgångsrikt med. En nyttjare av Wikipedia kan nu lyssna på innehållet och lära för livet. En detaljerad sammanställning av de målgrupper och behov vi identifierat och arbetat för finns redogjort i tabellen i Bilaga 1.

3 Redogörelse av projektresultatet

3.1 Beskriv kortfattat projektets mål och leveranser

Wikispeech-mjukvaran består av två distinkta delar som stärker varandra.

Den första delen är spelaren, en talsyntes (Wikispeech TTS), som är integrerade i MediaWiki-mjukvaran och som sådan har ett stort antal potentiella nyttjare och personer som är intresserad av att den ska bli bättre och förhoppningsvis kan tänka sig att bidra genom crowdsourcing.

Den andra delen är Talresursinsamlaren, vilket är en grupp med verktyg som gör att volontärer och andra intresserade kan bidra med taldata och förbättra existerande taldata.

Talresursinsamlaren länkas till från Talsyntesen för rättningar när något inte fungerat som förväntat. De rättningar som görs eller den nya data som samlas in används sedan för att göra Talsyntesen bättre över tid.

På grund av det inneboende beroendet lades en stor del resurser under projektet på att möjliggöra aktiveringen av Talsyntesen på Wikipedia. Då de krav och förutsättningar som fanns ändrades under projektets gång krävdes det ett flertal justeringar för att Talsyntesen skulle fungera, vilket tyvärr gjorde att Talresursinsamlaren inte helt kunde färdigställas och aktiveras inom ramen för projektet, även om mycket stora steg framåt togs. Dessa justeringar berodde på olika orsaker och godkändes kontinuerligt av PTS.

Mål och leveranser i ansökan:

1. Framtagande av ett system för att genom crowdsourcing samla inspelningar av tal på olika språk.

- Att samla in stora mängder taldata från en stor grupp människor, en grupp som behöver ha verklig mångfald, är en utmanande uppgift. Vår förhoppning är att crowdsourcing kan erbjuda en lösning.

Crowdsourcing fungerar dock endast om (1) medvetenhet kan nås och om (2) den medvetenheten kan omvandlas till engagemang och det (3) materialet sedan kan tas tillvara. Vi är övertygade om för att uppnå 1 och 2 behöver vi integrera verktygen för talresursinsamling i Talsyntesen.

Talsyntesen blir ingången från Wikipedia för de som vill bidra med talresurser. D.v.s. när någon exempelvis hör ett fel i hur ett ord uttalas kan hen bidra med en rättning, precis som att hen kan rätta en artikel på Wikipedia. Baserat på den här insikten har vi kontinuerligt omprioritera resurser till att säkerställa att Talsyntesen ska kunna aktiveras när nya oväntade utmaningar har uppkommit. Vi vill att när vi väl aktiverar Talresursinsamlaren ska allmänheten börja bidra för att Talsyntesen de använder ska bli bättre!

- I slutändan av projektet hade vi en färdig talsyntes, där lyssnaren själv kan förbättra resultatet. Talsyntesen har framgångsrikt genomgått en kod- och prestandagranskning av Wikimedia Foundation, och därmed är redo att aktiveras på Wikipedia. När så sker kommer Talsyntesen att vara integrerade i MediaWiki-mjukvaran som Wikipedia använder.

På grund av stora förändringar i Wikimedia Foundations ledarskikt de senaste månaderna, och därmed en ovilja att fatta beslut om långsiktiga åtaganden, kommer det dock att dröja en tid innan den integrationen sker. Därför har vi även utvecklat en s.k. *finess* (en *gadget*) som gör att Wikimedia Sverige kan drifva Wikispeech på egen serverinfrastruktur. För slutanvändaren är skillnaden ganska liten, men en finess är inte lika integrerad i mjukvaran utan kräver en separat server och att det skickas lite mer data fram och tillbaka. Därtill måste Wikimedia Sverige sköta en större del av underhållet över tid. I dagsläget bekostar Wikimedia Sverige kostnaden för servern med egna medel vilket innebär att om tjänsten blir väldigt populär och välanvänd kommer vi att behöva hitta finansiering för fortsatt drift. Kod för

tillägget finns här:

<https://github.com/wikimedia/mediawiki-extensions-Wikispeech>
h. För att testa Wikispeech behöver du skapa ett konto på svenskspråkiga Wikipedia, sedan finns instruktioner här: <https://sv.wikipedia.org/wiki/Wikipedia:Wikispeech>.

- Vi har under projektet byggt om arkitekturen för Talsyntesen och utvecklat dokumentationen för att det ska vara lättare att driftsätta Talsyntesen i andra miljöer, som exempel kan en myndighet mycket enklare aktivera Wikispeech på en wiki de själva använder som en del av sin IT-infrastruktur. Dokumentation (med länk till repot) finns här: https://www.mediawiki.org/wiki/Extension:Wikispeech/Installing_Speechoid.
- Grundstommen har byggts för en plattform för taldatainsamling samt annotering av taldata. Vi inledde arbetet med en analys av Mozilla Common Voice där vi identifierade styrkor och svagheter. De insikterna har vi sedan tagit vidare och vi har under projektet utvecklat en genomtänkt arkitektur med planer för hur metadata om den röst som är inspelad ska samlas in och hanteras, var vi sparar inspelningar, hur olika dialekter ska hanteras med mera. En annan viktig aspekt är hur det material som samlas in kan kvalitetssäkras av andra volontärer. Talresursinsamlarens repo finns här: <https://github.com/wikimedia/mediawiki-extensions-WikispeechSpeechDataCollector>.
- För Talresursinsamlaren har vi arbetat med att utveckla en manuskomponent. Detta är en viktig pusselbit då ett väl uttänkt manus gör det möjligt att begränsa mängden taldata som behöver samlas in för att skapa nya röster till Talsyntesen. Volontären guidas genom manuset till att spela in extra värdefulla meningar. Det innebär att det inte behöver spelas in lika mycket ljud och volontärinsatsen för att kunna göra Talsyntesen bättre minskar därmed. Repot finns här: <https://github.com/stts-se/wikispeech-manuscriptor>.
- Pandemisituationen har omöjliggjort delar av planerna runt tester, då det inte varit möjligt under rådande omständigheter att ta in försökspersoner i labb för att utveckla tester. Detta steg ströks för att istället ta fram, validera och reliabilitetstesta webbaserade metoder. Tre olika metoder har arbetats på. I slutändan fastslogs valideringsmetoden till en version av ARS ("Audience Response Systems") där de som testar talsyntesen klickar när något låter konstigt, så att det kan åtgärdas av vårt team senare.

2. Framtagande av komponenter som kan användas för att annotera det inspelade talet.

- STTS arbetade aktivt med att utveckla ett flertal centrala komponenter som kommer att göra det möjligt att annotera det inspelade talet. Det fanns dock inte möjlighet att under projektet färdigställa ett användargränssnitt. Insatsen för att ta fram en första version av detta är dock förhållandevis liten, men en större insats kommer att krävas för att säkerställa ett användarvänligt verktyg med ett minimum av buggar. Repot finns här: <https://github.com/stts-se/wikispeech-annotator>.
- I Talresursinsamlaren finns det stöd för att användare kan betygsätta varandras inspelningar. Ett gränssnitt behöver dock utvecklas.

3. Genomförande av evenemang för att finna metoder för att engagera allmänheten att bidra med inspelat tal och annoteringar samt ta fram ett färdigt kit för att andra enkelt ska kunna ordna egna talinsamlings/-annoteringsevenemang.

- Tillsammans med Dyslexiförbundet genomförde Wikimedia Sverige ett experimentellt evenemang med 16 deltagare. Genom evenemanget kunde vi testa det presentationsmaterial som tagits fram och få en bättre förståelse för utmaningarna och möjligheterna med att genomföra dylika evenemang.
- Baserat på resultatet från evenemanget togs därefter ett toolkit fram för att liknande evenemang snabbt ska kunna organiseras av olika Wikimediaorganisationer eller andra intressenter på olika platser i världen.
- En förbättrad projektportal har tagits fram. Informationen fokuserar på hur intresserade kan bidra med taldata. Portalen kan besökas här: <https://meta.wikimedia.org/wiki/Wikispeech>.
- Ett stort antal aktiva volontärer och organisationer inom Wikimediarörelsen har kontaktats om projektet för att skapa intresse och engagemang. Vår förhoppning är att det toolkit som tagits fram kommer att kunna återanvändas av dem för lokala evenemang i framtiden.

4. Utveckla system och strategier för att insamlat tal samt annoteringar skall komma Wikipedia och andra Wikimediaplattformarna till godo samt kunna samlas in via dessa plattformar.

- Ljudfiler har utgjort en försvinnande liten del av all media som finns på Wikimediaplattformarna och vi har därför behövt

undersöka vilka möjligheter och begränsningar som finns. Under projekttiden har vi därför genomfört ett antal olika aktiviteter där vi har fördjupat vår förståelse om hur ljudfiler bäst kan struktureras på Wikimedia Commons.

- Wikidata är en av Wikimediaplattformarna där länkad, strukturerad data lagras. Sedan 2018 har det även varit möjligt att integrera lexikala resurser på Wikidata. Att knyta samman dem med de lexikon vi har skulle skapa mervärde och ett ökat intresse. Det är särskilt intressant med tanke på att den gemenskap som redigerar lexikala resurser på Wikidata är relativt liten. Vår aktivitet på området skulle kunna leda till att fler ser värdet av lexikografiska data på Wikidata, både inom Wikimedia's befintliga volontärgemenskap och nya bidragsgivare med expertis inom språk och språkteknologi. Därför har vi undersökt möjligheterna för integration, även om det inte var aktuellt i detta läge då de lexikala resurserna ännu aktivt utvecklas på Wikidata. Då vi säkrade ett projektbidrag från Postkodstiftelsen kunde vi genomföra ett antal praktiska tester där resultatet är till nytta för projektet. En fallstudie finns här: https://meta.wikimedia.org/wiki/FindingGLAMs/White_Paper/WORD.
- Inom ramen för projektet har kontakter med samtliga Sveriges myndigheter skett. Målet har varit att skapa medvetenhet om projektet och utröna hur MediaWiki-mjukvaran används av dem och om de vore intresserade av att aktivera Wikispeech för ökad tillgänglighet. Förhoppningen är att över tid etablera ett nätverk av organisationer som använder Talsyntesen och därför väljer att aktivt bidra till att göra Wikispeech bättre på olika språk. Rapporten finns här: <https://se.wikimedia.org/wiki/Wikispeech/MediaWiki-inventering>.

5. Säkerställa att insamlat tal och annoteringar kan lagras på plattformar som möjliggör brett tillgängliggörande.

- Strukturer för intern lagring har utvecklats klart.
- KTH leder arbetet med Språkbanken Tal och vi har därför haft upprepade diskussioner om hur de framtida taldataresurserna bäst kan integreras där.
- Vi har haft ett antal möten med Mozilla Foundation om deras projekt Common Voice för att informera om projektet och eventuella framtida samarbeten.

Utöver de specifika projektmålen har vi även utbytt erfarenheter med Wikimedia Deutschland för att kunna utveckla Wikispeech till högsta möjliga kvalitet och integrera det bättre med Wikidata. En rapport från besöket finns här:

[https://se.wikimedia.org/wiki/Projekt:Wikispeech - Talresursinsamlingen 2019/WMDE site visit 2020-01-20-24](https://se.wikimedia.org/wiki/Projekt:Wikispeech_-_Talresursinsamlingen_2019/WMDE_site_visit_2020-01-20-24).

Efter att det blev uppenbart att vi inte skulle kunna aktivera Talsyntesen som en integrerad del av MediaWiki, och den tid som återstod inte skulle vara tillräcklig för att bygga färdigt Talresursinsamlaren, valde vi efter konsultation med PTS att revidera vårt arbete de sista månaderna av projektet. Vi lade resurser på två områden: 1. Att genomföra en inventering av svenska myndigheters nyttjande av MediaWiki, och deras intresse för Wikispeech samt 2. Att utveckla Wikispeech till en s.k. finess (istället för ett MediaWiki-tillägg), för aktivering under sommaren 2021.

3.2 Viktiga insikter

Vi hade redan från början insikten av behovet av att knyta Talresursinsamlaren till Talsyntesen. Att det var centralt för att kunna skapa engagemang för crowdsourcing hos Wikimediagemenskapen och att det är denna koppling som i deras ögon skiljer oss från andra taldataisamlade alternativ inom den större rörelsen för fri kunskap.

Vi förväntade oss en måttlig insats för att modernisera kodbasen samt åtgärda uppmärksammade problem med originalimplementeringen. Åtgärderna var dock större än förväntat då MediaWiki-mjukvaran under projektperioden utvecklades inom flera relevanta områden vilket krävde att vi byggde om Talsyntesen. Vi insåg tidigt att det skulle ta mer kalendertid att få spelaren driftsatt. Som ett resultat av detta justerade vi vår planering så att arbetet med spelaren intensifierades under de tidigare etapperna. Här såg vi vikten av att ha starka kontakter med Wikimedia Foundation och andra organisationer som aktivt bidrar med utveckling av MediaWiki för att få återkoppling, något vi nyttjade även för Talresursinsamlaren där vi redan under design-/arkitekturstadiet begärde återkoppling från våra kontakter.

En annan viktig insikt var de svårigheter det beroende vi hade gällande kod-/säkerhetsgranskningen av den mjukvara vi utvecklat. Även om vi på vår sida var färdiga med relevant utveckling gjorde det faktum att Wikimedia Foundation drog ut på tiden (vilket de förklarade berodde på covid-19 samt stora organisatoriska förändringar under 2020-2021) att viktig återkoppling senarelades och deras vilja och förmåga att driftsätta försvann.

Den komplexitet vi identifierade med att få Talsyntesen aktiverad på Wikipedia fick oss att inse att en implementering av Talresursinsamlaren borde innehålla vissa delar som integrerades i MediaWiki-mjukvaran, medan andra delar borde byggas som separata tjänster för att öka flexibiliteten. I de fall där det var lämpligt att bygga den som ett MediaWiki-tillägg kunde vi dra nytta av redan existerande funktionalitet inklusive den som byggts för Talsyntesen. Att köra den på en separat tjänst utanför Wikipedia innebar att vi kunde undvika den flaskhals som driftsättning av Wikimedia Foundation utgjorde. Samtidigt sker implementeringen på ett sådant sätt att den *kan* inkluderas direkt på Wikipedia vid ett senare tillfälle. Se följande dokument för analys:

[https://se.wikimedia.org/wiki/Projekt:Wikispeech - Talresursinsamlaren_2019/Integrering_av_Talresursinsamlaren](https://se.wikimedia.org/wiki/Projekt:Wikispeech_-_Talresursinsamlaren_2019/Integrering_av_Talresursinsamlaren).

4 Redogörelse av projektets genomförande

4.1 Genomförandeplan/tidplan

Projektets tidplan har i flera omgångar reviderats, i samråd med PTS, för att hantera förändringar av både extern natur (covid-19, förändringar hos Wikimedia Foundation) och intern natur (förseningar kring initial rekrytering, upptäckt behov av omimplementering av tidigare lösning). Längden på de individuella etapperna har ändrats vid flera tillfällen och slutdatumet för projektet har även förlängts 2021-03-31 till 2021-04-30.

Projektet blev inledningsvis försenat på grund av personalbrist. Detta berodde dels på en nyrekrytering som blev försenad, dels på andra projekt vilka blivit förlängda och därmed band upp teknisk personal längre än planerat.

En tidig insikt var att integrering av spelaren i Talsyntesen skulle kräva mer kalendertid (för att Wikimedia Foundation skulle kunna genomföra sina delar), vi justerade därför planeringen så att arbetet med Talsyntesen intensifierades under de tidigare etapperna och delar av arbetet med Talresursinsamlaren senarelades. Integration med Talsyntesen krävde även mer resurser än planerat. Vi insåg att MediaWiki hunnit förändrats mer än vad vi trott sedan vi byggde den första versionen. Två centrala implementationsval som gjorts i första versionen identifierades även som problematiska vilket krävde omstrukturering och omimplementering. En av dessa finns beskriven här:

<https://github.com/wikimedia/mediawiki-extensions-Wikispeech/blob>

[/master/docs/adr/0001-store-synthesized-audio-in-mediawiki-extensi](#)
[on.md](#). En sammanställning av alla finns i [WMDE-besöksrapporten](#):
[https://se.wikimedia.org/wiki/Projekt:Wikispeech - Talresursinsamla](https://se.wikimedia.org/wiki/Projekt:Wikispeech_-_Talresursinsamla)
[ren_2019/WMDE_site_visit_2020-01-20-24](#). Genom ombudgetering
efter minskade resekostnader, på grund av pandemin, bekostades
arbetstiden för detta.

Det har därefter varit en återkommande avvägning om att lägga energi
på Talsyntesen gentemot Talresursinsamlaren. På grund av den
föränderliga information vi fick från Wikimedia Foundation blev det
några turer fram och tillbaka, men vi valde i slutändan att omprioritera
så att Talsyntesen skulle vara så klar som möjligt inför den kod- och
prestandagranskning Wikimedia Foundation slutligen genomförde.

Covid-19 hade stor påverkan på vår planering (se sammanställningen
på: https://se.wikimedia.org/wiki/Fil:Justeringsbegäran_covid-19.pdf).

- Extern kommunikation (konferenser) ströks i princip helt under
en period och Resor och evenemang utgick helt. Eftersom det i
båda fall inte finns en möjlighet att ta igen dessa vid ett senare
tillfälle flyttades arbetstiden till andra aktiviteter (se nedan). De
icke-arbetstidskostnader som uteblev omvandlades till extra
arbetstid vilken användes för utveckling av Talsyntesen (se
nästkommade punkt).
- Extra kostnader kring projektledning tillkom som ett resultat av
pandemin.
- Ett behov av att designa om processen för användartester när
dessa inte längre var möjliga att genomföra på plats medförde att
denna komponent behövde extra resurser.
- Då vissa centrala evenemang ställts in krävdes andra
kommunikationsinsatser för att förankra våra planer med
nyckelspelare inom Wikimediagemenskapen.

I samband med att vi fick information om att Talsyntesen inte skulle
kunna aktiveras på Wikipedia ställde vi, i samråd med PTS, om så att all
utveckling och utvärdering av den centrala plattformen för
Talresursinsamlaren ställdes in och fokus istället lades på att
Talsyntesen. STTS fortsatta dock arbeta med utvecklingen av moduler
som är av generell nytta, även utanför Talresursinsamlaren. Parallellt
med detta ställdes kommunikationsinsatserna om från att vara riktade
mot Wikimediagemenskapen till att vara fokuserade på att identifiera
framtida brukare av Talsyntesen (myndigheter).

4.2 Projektets budget

Poster	Budget		Utfall		Andel av totala kostnader (%)
	PTS finansiering	Egenfinansiering	PTS finansiering	Egenfinansiering	
Projektledning	178 609 kr	0 kr	233 585 kr	0 kr	7,8 %
Community konsultation	178 775 kr	0 kr	212 888 kr	0 kr	7,1 %
Kommunikation	176 459 kr	70 000 kr	172 708 kr	70 000 kr	5,8 %
Integrering med spelare	379 301 kr	247 324 kr	479 106 kr	247 324 kr	16,0 %
Utveckling	861 250 kr	247 324 kr	987 111 kr	247 324 kr	32,9 %
Expertdiskussioner, rapportskrivande	224 766 kr	225 000 kr	205 893 kr	225 000 kr	6,9 %
Användartester	48 511 kr	0 kr	84 254 kr	0 kr	2,8 %
Arbetstid för referensgrupp, evenemang samt kommunikation	24 000 kr	0 kr	24 000 kr	0 kr	0,8 %
Resor och evenemang	385 625 kr	0 kr	51 661 kr	0 kr	1,7 %

Overheadkostnader	192 703 kr	290 000 kr	198 794 kr	290 000 kr	6,6 %
Lokaler	350 000 kr	0 kr	350 000 kr	0 kr	11,7 %
Summa	3 000 000 kr	1 079 648 kr	3 000 000 kr	1 079 648 kr	

Kostnader för underleverantörer.

STTS – Södermalms talteknologiservice AB		
Poster	Budget	Utfall
Utveckling	660 000 kr	662 500 kr
Integrering med spelare	150 000 kr	176 500 kr
Expertdiskussioner/rapportskrivande	140 000 kr	111 000 kr
Summa	950 000 kr	950 000 kr

KTH – Tal Musik och Hörsel (TMH) och Språkbanken Tal (SBTal)		
Poster	Budget	Utfall
Kommunikation	160 840 kr	114 971 kr
Utveckling	41 699 kr	41 699 kr
Användartester	48 511 kr	84 254 kr
Expertdiskussioner/rapportskrivande	309 766 kr	319 893 kr
Overheadkostnader	336 490 kr	336 490 kr
Summa	897 307 kr	897 307 kr

Dyslexiförbundet		
Poster	Budget	Utfall
Arbetstid för referensgrupp, evenemang samt kommunikation	24 000 kr	24 000 kr
Summa	24 000 kr	24 000 kr

4.3 Projektets arbetssätt

Wikimedia Sverige (WMSE) var projektägare och kontaktpart gentemot PTS under hela projektet. Både KTH och STTS rapporterade till Wikimedia Sverige som därefter rapporterade vidare till PTS.

KTH och STTS besitter omfattande expertis inom utveckling av talteknologi. Vi hade redan under förstudien utvecklat rutiner för hur samarbetet skulle komma att läggas upp och vilka styrkor och kompletterande kompetenser respektive team har.

WMSE har expertis på Wikimediaplattformarna, på crowdsourcing och har ett omfattande kontaktnät bland organisationer som arbetar med fri kunskap samt FOSS-mjukvara (vilka är viktiga mottagare av talresurserna som genereras).

KTH forskar inom området och bygger nu även upp Språkbanken Tal, vilket kommer att vara en viktig mottagare av materialet som detta projekt genererar i framtiden. Inom ramen för Språkbanken Tal kommer KTH bl.a. att arbeta med att ta fram en del verktyg som kan återanvändas i projektet, utreda det juridiska om hur taldata kan lagras, presentera samarbetet på olika konferenser och vetenskapliga artiklar och rapporter. Detta projekt kunde därför direkt dra nytta av detta utan att de kostnaderna behövde bäras av detta projekt.

STTS har expertis av att utveckla verktyg för att samla in talresurser och annotera dessa.

WMSE, KTH och STTS har arbetat tillsammans sedan 2015 med att göra Wikipedia mer tillgängligt. De olika teamen har kompletterande kunskap och erfarenhet och har god erfarenhet av att arbeta tillsammans. Internt har organisationernas team arbetat länge tillsammans, vilket gör att vi känner oss trygga i att arbetet kommer att gå smidigt.

Wikimedia Deutschland (Tyskland) är en ideell förening som är lokaliserad i Berlin. För att utbyta erfarenheter och inhämta kunskap arbetade två personer från Wikimedia Sverige under en vecka på Wikimedia Deutschlands kontor. Detta gav oss en möjlighet att bygga ett förtroende och personkännedom hos de olika teammedlemmarna. Wikimedia Deutschland har under projektet fungerat som bollplank och hjälpt oss att identifiera vägar framåt.

Wikimedia Foundation har huvudkontor i San Francisco och ansvarar för driften av Wikipedia och de andra Wikimediaprojekten. De

genomförde en kod- och prestandagranskning av den MediaWiki-mjukvara som vi utvecklat.

Dyslexiförbundet bidrog till projektet genom att delta i en referensgrupp för att testa mjukvaran samt med att hitta formen för hur inspelningsevenemang kan anordnas.

Följande organisationer fick vid olika tidpunkter ta del av information om projektet: WikiBlind (intresseorganisationen för blinda Wikipediaanvändare), Mozilla Foundation, Myndigheten för tillgängliga medier, Stockholm universitets tolkavdelning, Kungliga biblioteket, Uppsala universitetsbibliotek, Talkamatic, Karolinska institutet, Ludwig-Maximilians Universität München Institut für Phonetik und Sprachverarbeitung (IPS) och CLARIN ERIC.

5 Vad händer nu?

På kort sikt

Vi har mjuklanserat Wikispeech som en finess och kommer fortsätta att arbeta aktivt, med egen finansiering, för att förbättra användarupplevelsen baserat på den feedback vi mottar från de tidiga testarna.

Successivt kommer vi att kommunicera till större och större grupper och även genomföra tester på evenemang, när samhället öppnat upp igen efter pandemin. Vi räknar med att genomföra ett flertal uppföljningsprojekt och söker aktivt projektmedel för detta.

Ett formellt lanseringsevenemang för att öka synligheten för projektet kan komma att ske mot slutet av 2021, om pandemin tillåter fysiska evenemang. Företrädare från PTS kommer då att bjudas in.

1-2 år

Vi hoppas kunna aktivera Wikispeech som ett officiellt tillägg i MediaWiki som en beta-version. Då kommer antalet personer som testar mjukvaran att ha ökat markant och vi kommer att arbeta för att åtgärda problem de identifierar. Det kommer även att bli ett viktigt test för om mjukvaran kan skala till ett väldigt stort antal simultana användare.

Vi implementerar andra typer av röster m.m. som har utvecklats som är av högre kvalitet (d.v.s. att rösterna låter mer mänskliga).

Vi arbetar aktivt med testning av Talsyntesen på andra språk än svenska för att åtgärda konstigheter och specialfall.

Den grundläggande funktionaliteten för Talresursinsamlaren har färdigställts och vi har klarhet i vilka begränsningar GDPR skapar för vår insamling av taldata. Ett stort antal evenemang kommer att anordnas för att samla in taldata från olika underrepresenterade grupper.

Vi aktiverar de första verktygen som gör det möjligt att genom crowdsourcing aktivera nya språk. T.ex. ett segmenteringsverktyg där grundläggande information om hur meningar är uppbyggda på ett språk samlas in så att maskininlärning är möjlig att aktivera.

Wikidatas Lexem, som är strukturerad lexikografisk data, integreras med lexikonet i Wikispeech.

Ett flertal riktade aktiviteter för att utveckla lexikonet genomförs.

3-4 år

Vi kommer att etablera en arbetsgrupp med volontärer. Tillsammans med den kommer vi att föra etiska diskussioner om vilka röster som ska vara standard, vilka nya språk och röster som bör prioriteras, hur datainsamlingen kan nå så många underrepresenterade grupper som möjligt m.m.

Vi identifierar de komponenter som behövs för att aktivera nya språk med tillräcklig kvalitet och integrerar dem i Wikispeech. Om komponenterna är av bristfällig kvalitet eller helt saknas stödjer vi utvecklingen av dem. Om de är av god kvalitet med licenser begränsar användningen (t.ex. ett material en forskargrupp tagit fram) så arbetar vi för att de ska omlicensiera materialet.

Användarförbättringar aktiveras fullt ut. Det innebär att användaren kan rätta problem de stöter på (du hör ett ord som låter fel och vill rätta). De kan både skriva eller tala in hur det ska låta så att mjukvaran blir bättre överlag. Men det kommer även att vara möjligt att rätta ett fel direkt på sidan genom att märka upp att det ska uttalas på ett specifikt sätt i det sammanhanget.

Det kommer att kräva ett stort utvecklingsarbete, men om några år kommer Wikispeech att kunna aktiverat på de andra Wikimediaplattformarna, såsom Wikidata, Wikimedia Commons, Wikisource etc. Det ökar tillgängligheten på de plattformarna.

Wikispeech aktiverat på MediaWiki-wikis som andra organisationer använder som verktyg i sitt arbete. Vi har etablerat ett nätverk där organisationerna kan stödja arbetet på olika sätt (med utvecklartid eller annat arbete, med pengar eller annat).

Ett riktat arbete sker för att Wikispeech ska bli en pusselbit i kommunikationen i krissituationer där informationen behöver nå de som inte kan läsa. Till exempel behöver olika specialord kunna läggas till på kort sikt.

5-10 år

Insamlingen av taldata pågår parallellt på flera hundra språk för att Wikispeech ska kunna aktiveras. Ett stort antal samarbeten har formaliserats med olika aktörer över hela världen för att deras språkdata ska kunna användas i arbetet.

Wikispeech användningsområde breddas och de kan aktiveras på andra verktyg för att öka tillgängligheten generellt.

Specifika projekt och verktyg som syftar till språkbevarande eller insamlande av muntliga källor (*oral citations*) har aktiverats, med grund i Talresursinsamlaren.

Bilaga 1. Målgruppens behov och hinder

Fyll i identifierade behov och hinder för respektive målgrupp samt vilken åtgärd ni identifierat för de olika hindren. I kolumnen "Status" beskriver ni hur långt ni kommit i att åtgärda hindret.

Målgrupp	Identifierat behov	Identifierat hinder	Identifierad åtgärd	Status
Wikimedia användarna	Wikimediaplattformarna är exempel på kunskapsspridare. För att effektivt kunna sprida kunskap behöver den presenteras på sätt som gör det möjligt för alla konsumenter att tillgodogöra sig informationen. En stor del av världens befolkning kan inte läsa eller lär sig bättre av att lyssna.	Att kunna ta del av materialet genom en välfungerande talsyntes förbättrar lärandet, men det finns i dagsläget inte någon fritt licensierad talsyntes som kan integreras på Wikipedia. Att kunna samla in talresurser i stor skala möjliggör för en bättre talsyntes på bl.a. svenska men även aktivering på de övriga 300 språken av Wikipedia.	Arbete för att aktivera talsyntesen på de svensk-, engelsk- och arabiskspråkiga Wikipediaversionerna och vidareutveckling baserat på återkopplingen som ges.	Talsyntesen fungerar på svenska, engelska och arabiska. Efter projektets officiella slut har vi även aktiverat den på norska och baskiska. Ett par feedbackrundor har genomförts och vi har fortsatt vidareutveckla talsyntesen i oförändrad hastighet till och med juli 2021.
Forskare	För att bedriva forskning runt tal behövs i många fall stora mängder tillgänglig och annoterad taldata.	För forskning runt svenska och andra språk saknas data för många typer av forskningsprojekt helt eller delvis. Det saknas fritt licensierat och väl strukturerat och dokumenterat material att använda.	Vi har diskuterat vilka databehov olika forskargrupper har och vad vi kan göra för insatser via crowdsourcing.	Innan vi har ett aktivt insamlingsarbete av språkdata är samarbetena och stödet begränsade till samtal om möjligheter.

<p>Språkbevarare</p>	<p>Att kunna samla in inspelningar och annoteringar av tal hjälper även till att bevara olika språk. Inte minst hjälper välfungerande tjänster på det egna språket (där Wikimediaplattformarna är centrala) att skapa intresse från ungdomar att lära sig språket och nyttja det. Talresursinsamlaren kan även hjälpa till att samla in taldata för att bevara språket.</p>	<p>Verktyg som fungerar på hotade språk är ovanliga då de inte prioriteras av kommersiella aktörer på grund av sin storlek, och det blir en ond cirkel.</p> <p>Det saknas delvis även välfungerande verktyg samt metodik för att effektivt samla in taldata för språkbevarande i den takt som behövs.</p>	<p>Vi har presenterat om projektet på konferensen Arctic Knot 2021, som fokuserade på bevarande av minoritetsspråk, samt jobbat med Baskiska Wikipedia för att kunna aktivera Wikispeech där som en liten pilot.</p>	<p>Wikispeech på Baskiska Wikipedia är aktiverat, men vi har inte kunna arbeta för att lösa problem där ännu. Innan vi har ett aktivt insamlingsarbete av språkdata är samarbetena begränsade till samtal om möjligheter.</p>
<p>Återanvändare av verktyget</p>	<p>Talresursinsamlaren är skapad för MediaWiki, en mjukvara som används av tusentals olika organisationer. Det vi utvecklar för ett visst behov på Wikimediaplattformarna kan därför komma att nyttjas för ett helt annat användningsområde (antingen som helhet eller så kan vissa specifika delar vara relevanta).</p>	<p>För att detta ska vara möjligt krävs att så många delar som möjligt av verktyget görs tillgängligt för återanvändning. Att fritt licensiera allt material möjliggör spridning.</p>	<p>Allt material som skapats är fritt licensierat och vi har lagt mycket resurser på att förbättra dokumentationen.</p>	<p>I dagsläget har vi inte öppnat upp vår server för andra organisationer eller privatpersoner att använda verktyget. Däremot kan de förstås ta en kopia av allt och sätta upp verktyget på sin egen IT-infrastruktur då det är fritt licensierat.</p>
<p>Språkstudenter</p>	<p>En välfylld ordlista med korrekt uttal av ord är värdefullt för att lära sig ett språk. Vilka ord som är centrala varierar dock beroende på var i utbildningskurvan en befinner sig.</p>	<p>Att skapa eller integrera ordlistor är tidskrävande. Våra ordlistor kan antingen utvecklas genom att vi inkluderar existerande ordlistor från olika aktörer, eller skapar nya ordlistor med</p>	<p>Det finns många existerande ordlistor som vi kan komma att ladda upp, men det kräver en hel del arbete.</p> <p>Volontärer kan engageras i arbetet via Wikidata, med de</p>	<p>Wikidatas lexikala resurser fortsätter att utvecklas av en liten grupp dedikerade volontärer.</p> <p>Vi har fört samtal om ordlistor kan släppas med</p>

		hjälp av crowdsourcing, eller en kombination av de två.	lexikala delarna av Wikidata är ännu under utveckling. Ett aktivt arbete för att skapa engagemang behöver ske.	flera olika myndigheter och andra aktörer.
Myndigheter	Myndigheter arbetar under strikta krav på tillgänglighet, där talteknologi är en av de lösningar som används.	Då myndighetsmaterial ofta innehåller specifikt språk behövs specialordlistor för att talteknologin ska fungera på bästa sätt. Många myndigheter känner inte till möjligheterna med wikis som verktyg för deras arbete och inte heller att Wikispeech kan göra MediaWiki-wikis tillgänglighetsanpassade.	Att ha möjligheten att snabbt och kostnadsfritt få specialord adderade till en lexikal resurs kan spara stora summor och förbättra den tjänst de erbjuder. Vi behöver kontakta myndigheterna för att informera/utbilda dem om möjligheterna.	Vi har kontaktat alla svenska myndigheter om deras intresse av att använda verktyget. Resultatet har sammanställts i en rapport. Vi planerar att erbjuda dem utbildningar om MediaWiki och Wikispeech när verktyget blivit mer stabilt.
Tillgänglighetstjänster	För dessa tjänster behövs mer data för ytterligare utveckling. Därtill kan nya smarta lösningar som tas fram inom projektet återanvändas.	Det är kostsamt och arbetsintensivt att bygga upp språkresurser som kan användas.	Att nyttja crowdsourcing har potential att göra detta möjligt. Tjänsterna kan lära från våra experiment och själva börja samla in språkresurser via crowdsourcing. Gör de det med fria licenser gagnas alla.	Innan vi har ett aktivt insamlingsarbete av språkdata är samarbetena begränsade till samtal om möjligheter.

Talteknolog iutvecklare	För att skapa talteknologiska produkter för svenska, exempelvis tal-till-text, krävs mängder av inspelningar av svenskt tal med tillhörande annoteringar. Det krävs även flera olika fungerande mjukvarukomponenter.	Att samla in stora mängder taldata är svårt och alltför kostsamt för små aktörer. Det finns många komponenter som är fritt licensierade som kan användas, men på många språk saknas delar eller allt.	Förutom själva datan, kan de nya smarta lösningar som tagits fram inom projektet återanvändas, eftersom det publiceras som fri mjukvara.	Innan vi har ett aktivt insamlingsarbete av språkdata är samarbetena begränsade till samtal om möjligheter.
Utvecklare med behov av tal (men ej talteknologer)	Talteknologi utgör en del av många andra discipliner, där utvecklare/verksamheter har behov av taldata för olika tillämpningsområden.	Det saknas fritt licensierad mjukvara.	Den mjukvara som byggs är fritt licensierad, väldokumenterad och heltäckande. Kommunikation om dess existens för att öka användningen är viktigt.	Mjukvaran är fritt licensierad och kan återanvändas av alla intresserade.
Enskilda individer	På individnivå vet vi att en mycket stor mängd människor finner tillfredsställelse i att bidra till t.ex. Wikipedia. Vi vet också att språk engagerar, och tror att den möjlighet till att delta i en verksamhet som formar talteknologin kan skapa en stark gemenskap och delaktighetskänsla hos många. Verktuget är en variation på det som i litteraturen brukar kallas "citizen science", som bidrar till att skapa en meningsfull förströelse för individer.	För att enskilda ska engageras behöver de ha medvetenhet om att de kan hjälpa till och stöd att komma igång.	Det behöver tas fram detaljerad dokumentation samt läro- och kommunikationsmaterial.	Vi har arbetat aktivt med att ta fram detaljerad dokumentation samt centrala delar av det läro- och kommunikationsmaterial som behövs. Materialet kan färdigställas när Talresursinsamlaren är komplett.