

Stockholm 2018-03-09

Ver 0.6

Slutrapport projektgenomförande - Wikimedia Sverige

8 Innehållsförteckning

<u>1</u>	<u>Inledning</u>	1
<u>2</u>	<u>Allmän information</u>	1
<u>3</u>	<u>Sammanfattning</u>	1
<u>4</u>	<u>Bakgrund</u>	1
<u>5</u>	<u>Syfte och målgrupp</u>	1
<u>6</u>	<u>Projektets mål</u>	1
<u>7</u>	<u>Redogörelse av projektresultatet</u>	2
<u>7.1</u>	<u>Beskriv kortfattat projektets resultat samt leverabler</u>	2
<u>7.2</u>	<u>Viktiga insikter</u>	2
<u>8</u>	<u>Redogörelse av projektets genomförande</u>	2
<u>8.1</u>	<u>Genomförandeplan/tidplan</u>	2
<u>8.2</u>	<u>Projektets budget</u>	2
<u>8.3</u>	<u>Projektets arbetsätt</u>	2
<u>9</u>	<u>Vad händer nu?</u>	2

1 Inledning

Notera att slutrapporten är en leverabel till PTS och den är en redovisning av hur projektet har gått, vilka resultat som uppnåtts och de slutsatser som gjorts. Slutrapporten bör inte överstiga 20 sidor.

Den dokumentation som ni angivit i förstudien att ni ska producera i samband med projektet ska ses som fristående från detta dokument.

2 Allmän information

Företag/Organisation:	Wikimedia Sverige, STTS, KTH
Projektnamn:	Wikispeech
Författare av detta dokument:	John Andersson, i samarbete med André Costa, Sebastian Berlin från Wikimedia Sverige, samt Nikolaj Lindberg, Hanna Lindgren, Harald Berthelsen från STTS, samt Joakim Gustafson, Jens Edlund, Jonas Beskow från KTH
Projektets tidplan (från datum – till datum)	Projektstart 2016-03-15. Projekt slut 2017-12-31.

3 Sammanfattning

Ange positiva effekter och därefter övriga erfarenheter av projektets genomförande samt kortfattad sammanfattning av resultatet. Bör vara ca 1-2 sidor långt kapitel.

Utvecklingen av Wikispeech pågick mellan mars 2016 och december 2017. Arbetet baserades på en förstudie som utarbetats av Wikimedia Sverige (WMSE) tillsammans med Södermalms Talteknologiservice (STTS) samt Kungliga tekniska högskolan (KTH).

Syftet med projektet var att skapa en fritt licensierad talsyntes som kan användas på Wikipedia och kontinuerligt utvecklas av användarna genom crowdsourcing. Ett innovativt koncept för att över tid kunna öka kvaliteten på talsynteser på olika språk, såväl som att kunna skapa nya talsynteser på fler språk. Vi bedömer att detta kommer att ha en mycket positiv effekt på vilka som kan tillgodogöra sig informationen. Inte minst när Wikispeech aktiveras på fler

språk där många av modersmålstalarna har en begränsad eller icke-existerande läsförståelse p.g.a. begränsad utbildningsnivå.

Under projektet låg fokuset initialt på att utveckla talsyntesfunktionaliteten i linje med den förstudie som tagits fram. Målet var att Wikispeech då skulle fungera på svenska, engelska och arabiska, vilket uppnåddes. Detta inkluderade spelare, lexikon och syntesmotor. Som en del av detta genomförde vi flera utvärderingar med representanter från intresseorganisationer för personer med funktionsnedsättning som kom med värdefull feedback. Vi bjöd in personer från dem att testa de tidiga versionerna av Wikispeech och ge oss feedback på den existerande funktionaliteten samt vilken ytterligare funktionalitet de såg behov av. Vi genomförde först ett tvåtimmarsmöte med en representant från Dyslexiförbundet samt därefter ett tvåtimmarsmöte med företrädare från Synskadades riksförbund. Då kunde de förevisa vad de sett som positivt och vad som kunde utvecklas ytterligare, visa på bra exempel på andra lösningar samt rekommendera prioriteringar för kommande utveckling.

KTH genomförde även tester med studenter rörande bl.a. hur effektivt en lyssnare tillgodogör sig information i längre texter. Ökad förståelse runt detta har en direkt effekt på vidareutvecklingen av Wikispeech (där fokuset är att säkerställa kunskapsöverföring via Wikipedia) men även för andra externa plattformar (ex. myndighetsinformation). En positiv effekt med våra diskussioner var att för arabiska släpptes en ny språkresurs för arabiska (en hts-röst) under en fri licens.

När detta var uppnått, efter ungefär $\frac{2}{3}$ av projekttiden passerat, flyttades fokus till rättningsfunktionaliteten. Det vill säga möjligheten för användarna att tillsammans utveckla Wikispeechs lexikon för att orden skall återges på ett korrektare sätt samt att underlätta möjligheten att kunna lägga till fler språk när språkresurser identifierats. Den tekniska lösningen för att användarna skall kunna bidra till lexikonet är på plats, men det återstår att färdigställa den användarvänliga editor som vi ser som nästa steg.

Vi identifierade flera intressanta möjligheter för dem att kunna bidra på ytterligare sätt och för olika syften som vi fortsätter undersöka. Detta inkluderar exempelvis inspelningar för att utveckla en tal-till-text-lösning samt att snabbt förbättra lexikonet i samband med en kris. De fritt licensierade lexikala resurser som med tiden kommer att vara ett resultat av projektet kommer att ha en positiv effekt på talsyntesutvecklingen på bl.a. svenska. Det kommer även att möjliggöra att fler aktörer kommer att kunna ta fram talsynteser, inte minst fritt licensierade lösningar som inte spårar användarna.

Kontinuerligt under hela projektet bedrevs ett mycket aktivt kommunikationsarbete för att säkerställa en förståelse för frågan och behovet av projektet hos både den breda allmänheten och forskare, men framförallt hos Wikimedia-rörelsen. Detta för att en aktivering av Wikispeech med tiden skall vara möjlig. Arbetet med tillgänglighetsanpassning av Wikimedia-projekten har delvis varit eftersatt och det bedömdes finnas ett behov av att informera om

behoven och möjligheterna för att entusiasmera relevanta aktörer. Genom projektet har vi kunnat sätta fokus på detta viktiga område.

För Wikimedia Sverige var detta det första större utvecklingsprojekt för att bygga ett tillägg till MediaWiki-mjukvaran som vi genomfört. Genom projektet kunde vi utveckla rutiner och bygga upp en kunskapsbas som kommer att vara mycket värdefull för kommande projekt.

Fokuset på att öka tillgängligheten för webbplatserna har även möjliggjort att vi i Sverige kunnat bygga upp central kompetens som även kan gagna andra utvecklingsprojekt inom talsyntesområdet. Det säkerställer även att fler språk som är prioriterade av svenska aktörer kan inkluderas. Vi har bl.a. skickat in en ansökan för att kunna inkludera minoritetsspråken i Wikispeech. Genom projektet har vi kunnat bygga upp ett nätverk vilket möjliggör nya framtida spännande projekt.

4 Bakgrund

Redogör kortfattat för bakgrunden till projektet.

Wikipedia (WP) är en av världens mest använda webbplatser med ca 500 miljoner besökare och ca 20 miljarder sidvisningar varje månad.

WP är en s.k. wiki och använder mjukvaran MediaWiki i bakgrunden. MediaWiki används av många tusen andra webbplatser och detta projekt syftar till att skapa den programvara som behövs för att talsyntes skall kunna användas på alla dessa men optimerad för WP. Artiklarna innehåller många fackuttryck vilket gör att talsyntesens uttalslexikon måste vara mycket omfattande. Därtill finns WP på strax under 300 olika språk, och plattformen skall vara skalbar till alla de språken, samt alla framtida.

Flexibilitet är därför centralt. Initialt beslutades att förbereda Wikispeech på svenska, engelska samt ett höger-till-vänster-språk (arabiska). All programvara och allt material som utvecklats är open source för att utvecklare från hela världen skall kunna hjälpa till. Utvecklingen av Wikispeech bygger på crowdsourcing-principer där vi på så sätt kan tillgodogöra sig den språkliga expertis som finns hos de tiotusentals volontärer som är involverade i Wikimedia's olika projekt. Crowdsourcing är den metod som alla projekten bygger på, och detta var därför ett naturligt val även för talsyntesen. Genom att användargenerera talsyntesen med inspelningar av de specialiserade texterna och utveckla lexikonet med uttalsbeskrivningar kan vi med tiden nå en förfinad och högkvalitativ talsyntes även i obskyra ämnen på språk som tidigare helt saknat en fungerande talsyntes. Detta projekt möjliggjorde uppsättningen av det ramverk som behövs för detta.

Allt producerat material, oavsett typ, är fritt licensierat och därmed vara möjligt att gratis återanvändas av vem som helst (inklusive kommersiella verksamheter). Fri licensiering är en central del av Wikimedia-rörelsens sätt att arbeta då andra skall kunna fortsätta och bygga vidare på arbetet.

Nya MediaWiki-tillägg driftas på Wikimedia Foundations servrar vilket säkerställer att arbetet lever kvar även efter projektslutet.

5 Syfte och målgrupp

Redogör för syftet med projektet och till vilken målgrupp projektet riktades.

Redogör för hur projektet har uppfyllt detta.

Syftet med projektet var att ta fram en fritt licensierad talsyntes anpassad för att med tiden kunna aktiveras på Wikipedias alla språkversioner.

Målgruppen var alla de läsare som lär sig bättre genom att lyssna eller av olika skäl inte kan läsa. Med ett fokus på personer med dyslexi och/eller synnedsättningar. Då behoven är olika för olika individer är möjligheten med egna personliga inställningar viktigt.

Som planerat finns en fungerande talsyntes på svenska, engelska och arabiska anpassad för Wikipedia. Vidareutveckling, baserat på inkommen feedback, kommer att fortsätta ske med förhoppning att full aktivering på Wikipedia sker under 2018-2019.

6 Projektets mål

Ange projektets mål (effektmål och resultatmål som specificerade i förstudiematerialet). Lämna redogörelsen i tabellform om det är möjligt.

Övergripande leverabler

Leverabel	Uppfyllnad
En talsyntes som ett tillägg (en. <i>extension</i>) för MediaWiki (mjukvaran som driver Wikipedia) byggd på öppen programvara som inte kräver en lokal installation hos slutanvändaren (detta sköts på serversidan).	En serverbaserad extension har utvecklats, med majoriteten av de ingående delarna som identifierades i förstudien. Den funktionalitet som inte gick att bygga under projektet var variering av uppspelningshastighet (vilket är högprioriterat och kommer att byggas under 2018 för andra medel), ljudnotifikationer för ex. länkar och fotnoter vilket prioriterades ned efter diskussioner med intresseorganisationer för personer med funktionsnedsättning, samt uppspelning av informationen på den omgivande sidan runt brödtexten (menyer, knappar, sidfot m.m.) vilket prioriterades ned för att istället säkerställa högre kvalitet på brödtexten då intresseorganisationerna rekommenderat detta.

	För rättningen visade det sig orimligt att i dagsläget lägga till rättningar i enskilda artiklar då Wikimedia Foundation ännu inte aktiverat den underliggande funktionalitet som behövs för att rättningarna skall kunna behållas när artiklarna ständigt redigeras och ändras. Istället fokuserade vi på rättning och utökning av lexikonet.
En flexibel och väldokumenterad struktur för fortsatt vidareutveckling finns på plats.	Ett omfattande arbete har lagts på dokumentation av programvaran, av automatiserade tester, av det arbete som skett i projektet, samt om olika behov som kan utvecklas i kommande insatser.
Tre språk förberedda som proof of concept, varav ett höger-till-vänster-språk.	Svenska, engelska och arabiska (höger-till-vänster-språk med icke-latinska alfabet) har förberetts som proof of concept. Därtill har visst arbete skett även för spanska och norska.

Effektmålen

Målgrupp	Effektmål	Mätmetod vid uppföljning	Uppfyllnad
Wikipedia-användare som lär sig bättre av att lyssna är ca. 25% av användarna. Det totala antalet Wikipedia-användare är en halv miljard, d.v.s. 125 000 000 personer.	Talsyntesen används minst 10 000 gånger under den <i>sista projektmånaden</i> . (Antalet personer spåras inte då vi vill värna om integriteten.)	Uppföljning av hur mycket som tjänsten används genom att titta på loggar.	Under projektets gång ändrades Wikimedia Foundations rutiner, vilket tidigare rapporterats. Wikispeech finns i dagsläget aktiverat på Wikimedia Cloud Services. ¹ Det är en teknisk testmiljö med mycket mindre användning, vilket gör

¹ https://wikitech.wikimedia.org/wiki/Help:Cloud_Services_Introduction, hur Wikispeech används beskrivs här: <https://www.mediawiki.org/wiki/Extension:Wikispeech>

			detta mätvärde obsolet.
Svenska personer med lässvårigheter av olika orsaker, mellan 5-8% av befolkningen vilket innebär ca. 750 000 personer.	Efter aktivering på svenskspråkiga Wikipedia ska i en kontrollgrupp representativ för målgruppen en majoritet anse att talsyntesen underlättar informationsinlämningen för dem.	Både genom webbaserade tester och via kontrollerade användartester genomförda av experter på KTH.	Aktivering på svenskspråkiga Wikipedia bedöms ligga minst ett-två år bort i tiden. Testmetoder för syntes av längre sekvenser av informationsrikt tal saknas idag. Experter på KTH har tagit fram förslag på sådana metoder, och de har testats i kontrollerade försök.

7 Redogörelse av projektresultatet

I detta kapitel vill vi att du lägger fokus på själva projektresultatet. Dvs. vad blev slutsatserna av projektet, oavsett om det är ett fälttest, teknisk utveckling, en studie eller något annat som utförts. Lämna gärna rekommendationer som andra projekt kan ha nytta av.

7.1 Beskriv kortfattat projektets resultat samt leverabler

Här redogör du för projektets resultat. Det kan vara en kortfattat beskrivning t.ex. av hur många personer som har deltagit i en specifik undersökning och resultatet av undersökningen, det kan vara en beskrivning av en teknisk lösning och hur den ska användas etc.

Den fritt licensierade talsyntesen Wikispeech har utvecklats under projektet. Detta för att i första hand bidra med ökad tillgänglighet på Wikipedia, världens femte mest populära webbplats, för personer med lässvårigheter.

Utvecklingen har varit framgångsrik och en fungerande spelare med röster och lexikon finns nu för svenska, engelska och arabiska. Spelaren består av två komponenter. En del tolkar sidans innehåll och levererar denna till syntesmotorn den andra levererar ljudet till användaren och hanterar användarens interaktioner med sidan. Den feedback som inkom under projektet, samt att vissa delar vi identifierat som återanvändbara under förstudien behövde byggas om, ledde till att vi fick skjuta viss planerad funktionalitet på

framtiden. Vad som återstår att utveckla har noggrant dokumenterats i vårt *task management system, Phabricator*². Därtill finns grundläggande funktionalitet för att vidareutveckla kvaliteten på talsyntesen genom crowdsourcing.

För att säkerställa att Wikispeech uppfyller de behov som fanns från intresseorganisationerna för personer med funktionsnedsättning genomfördes två stycken tester på plats hos Wikimedia Sverige. Fokus under dessa tester var på hur Wikispeech kan användas om man har dyslexi eller en synskada.

Under 2018 är förhoppningen att Wikispeech skall kunna aktiveras för betatestning på de tre språkversionerna av Wikipedia.

För att få stöd för implementering och engagemang i det kommande crowdsourcingsarbetet har vi fokuserat mycket på kommunikation till olika intressenter och de många medskapare inom Wikimedia samt inom forskningssektorn. Vi har deltagit på flera konferenser och presenterat projektet samt har varit aktivt engagerade i organiseringen av *Wikimedia Diversity Conference 2017* för att lyfta behovet av den här typen av frågor. Där deltog 80 ledargestalter i Wikimedia-rörelsen från över 60 länder.

För att säkerställa att lyssnaren kan tillgodogöra sig artikelns innehåll och lära sig av materialet har KTH genomfört tester med studenter. Två metoder har tagits fram: ARS (Audience Response System) bygger på de screenings Hollywoods filmindustri använder sig av, och avser att peka ut problem i längre, kontinuerlig syntes. CCT (Continuous Comprehension Testing) är en extension av traditionella hörförståelsetester, där man lägger vikt vid att det är en längre, sammanhängande text som presenteras, som lyssnaren har för avsikt att tillgodogöra sig. Metoden ger även en bild av lyssnartrötthet, dvs. hur pass ansträngande det är att tillgodogöra sig innehållet under en längre tid. Båda metoderna har verifierats i initiala användartester med studenter. Bristen på effektiva utvärderingsmetoder för talsyntes och dess hämmande effekt på talsyntesutveckling är välkänd. De positiva resultaten av de inledande försöken med ARS och CCT har rönt brett intresse, och har lett till att flera ansökningar om konsolidering och utveckling av dessa metoder är under utarbetning, bland annat tillsammans med Myndigheten för Tillgängliga Medier, som arbetar med den typ av talsyntes som metoderna fokuserar på. Avsikten är bland annat att göra metoderna fritt tillgängliga via den nystartade Språkbanken Tal på KTH.³

Projektet har rönt mycket uppmärksamhet, och har omskrivits i nationell och internationell media 61 gånger.⁴ Det presenterades 2016 i en artikel på den prestigefyllda *Workshop on Speech Synthesis*, som går vart fjärde år och samlar

² <https://phabricator.wikimedia.org/project/view/1663/>

³ <http://sprakbanken.speech.kth.se/>

⁴ https://se.wikimedia.org/wiki/Projekt:Wikispeech_2016/Omnämmanden

hela den ledande industrin och akademien inom talsyntes,⁵ och arbete runt projektet har lett till ytterligare publikationer.⁶

7.2 Viktiga insikter

Nu när projektet har ett resultat så finns det alltid viktiga insikter som gjorts, det kan t.ex. vara att man insett att olika användargrupper har helt skilda behov, något som man kanske inte insåg från början, att man skulle utvecklat på en annan plattform, etc.

Viktiga insikter från detta arbete har varit den stora insats som krävs för att, i ett decentraliserat öppen programvaruprojekt som MediaWiki, få till stånd säkerhets- och kodgranskning. Denna granskning är ett krav när ny funktionalitet ska aktiveras i produktionsmiljö. Det har inte funnits etablerade rutiner bland Wikimedia-utvecklarna för att hantera större externa utvecklingsprojekt likt detta på ett effektivt och fullgott sätt. Det innebär att den kodgranskning vi hoppats få av Wikimedia Foundations utvecklare påbörjades först i december 2017, precis i slutet av projektet. De kommentarer vi fick då kommer vi att åtgärda under första halvan av 2018 och därefter arbeta för att ytterligare kod- och säkerhetsgranskning genomförs. De krav som detta innebär kommer vi att ta höjd för i kommande utvecklingsprojekt.

Det var ambitiöst att även inkludera rättning inom ramen för projektet då det visade sig att de komponenter som identifierats för spelaren vid närmare granskning inte uppfyllde våra behov. Bristerna hos dessa komponenter var dels relaterade till att de la belastningen på användarens dator, istället för våra servrar, dels att de enbart hanterade rå text, snarare än rik text, med rubriker och länkar. Vi tvingades även revidera vår ursprungliga plan att utgå från den syntax som används på Wikipedia för att istället utgå från den HTML som visas på sidan. Denna revidering innebär å andra sidan att Wikispeech lättare skulle kunna anpassas för användning utanför MediaWiki..

En bättre strategi skulle ha varit att inkludera rättningsfunktionalitet som stretch-goals istället då det var ett väl avgränsat separerat paket och nu tog tid och fokus från ytterligare förbättring av spelarens funktionalitet.

Vi såg även ett stort värde av att inkludera intresseorganisationer för personer med funktionsnedsättning i arbetet med att utveckla Wikispeech. Formatet för hur detta bäst organiserats skulle dock med fördel kunna utvecklas i ett separat projekt tillsammans med dem. Nu skedde det hela ad-hoc och det fanns inget material med exempelvis förslag på frågor att ställa dem som vi kunde ta del av.

Den tidsåtgång som krävs för att kunna hantera oförutsedda händelser och problem var en viktig lärdom. Trots en noggrann förstudie är det inte möjligt att

⁵ Andersson, J., Berlin, S., Costa, A., Berthelsen, H., Lindgren, H., Lindberg, N., Beskow, J., Edlund, J. & Gustafson, J. (2016). WikiSpeech—enabling open source text-to-speech for Wikipedia. In 9th ISCA Workshop on speech synthesis, Sunnyvale, CA, USA.

(http://ssw9.net/papers/ssw9_PS1-12_Andersson.pdf)

⁶ Malisz, Z., Berthelsen, H., Beskow, J., & Gustafson, J. (2017). Controlling prominence realisation in parametric DNN-based speech synthesis. In Proc. Interspeech 2017.

(http://www.isca-speech.org/archive/Interspeech_2017/pdfs/1355.PDF)

gå igenom alla delar och det räcker att någon del inte kan inkluderas för att en väldigt aktiv omstrukturering och planering skall krävas. Även om vi utvecklat iterativt var marginalerna för detta inte tillräckligt väl tilltaget i budgeten. Vi bedömer att upplägget med etapprapporter därför är väl lämpad då ändringarna kontinuerligt kan presenteras och godkännas av ansvarig handläggare på PTS.

8 Redogörelse av projektets genomförande

I detta kapitel vill vi ha en redogörelse av hur själva genomförandet av projektet har gått (vad har gått bra och vad har gått mindre bra och förslag på hur man kunde ha undvikit de värsta fallgroparna om man skulle göra om projektet – dvs. erfarenheter och insikter).

8.1 Genomförandeplan/tidplan

Lämna en kort redogörelse för de aktiviteter som du har genomfört i projektet och när under projekttiden de har ägt rum. Jämför med din ursprungliga projektplan och kommentera avvikelser från planen.

Vi inledde projektet med att sätta upp en struktur för samarbetet genom att använda task management-systemet Phabricator. Där lade vi in alla de uppgifter vi identifierat, inklusive de vi inte planerat att arbeta på inom ramen för projektet då vi ville säkerställa att dessa inte glömdes bort samt för att inspirera andra utvecklare till att bidra. Detta visade sig vara framgångsrikt då det fungerade väl för vår interna koordinering och även skapade intresse från andra utvecklare som stötte på projektet.

Under förstudien hade vi identifierat en mängd funktioner vi ville inkludera. Dessa prioriterades och vi hoppades kunna utveckla så många av dem som möjligt. Vi hann med en stor del av den tänkta utvecklingen men begränsades av tre faktorer:

1. En del av den existerande mjukvara vi identifierat var av lägre kvalitet än vad vi trott (bristande dokumentation, buggar m.m.) och kunde inte inkluderas utan vi var tvungna att utveckla egna lösningar.
2. Utvecklingen av MediaWiki var inom vissa områden kraftigt försenad varför viss funktionalitet inte var möjlig att bygga.
3. Andra delar av MediaWiki genomgick större förändringar under projekttiden. Detta innebar att vi var tvungna att gå tillbaka och åtgärda problem som uppkommit för oss när MediaWiki förändrats.

Det huvudsakliga fokuset låg därför på att utveckla spelaren och bakomliggande system (lexikon m.m.). Spelaren fungerar bra på både höger-till-vänster- och vänster-till-höger-språk och markerar både det ord och den mening som läses upp. Det är möjligt att hoppa fram och tillbaka i texten och även att markera specifika delar och spela upp dem. Funktionalitet har förberetts för användaren att göra egna inställningar. Av ovan angivna skäl tog det hela dock längre tid än ursprungligen planerat.

Rättningsfunktionaliteten utvecklades först under Etapp 5, och hann inte utvecklas och testas så mycket som vi hoppats. Det finns nu en prototyp för att lägga till förslag på tillägg till lexikonet.

Kommunikationen pågick under hela projektet och var framgångsrik. Detta var något vi såg som viktigt för Wikispeechs implementering. I Wikimedia-rörelsen finns det ett stort intresse för arbete runt tillgängliggörande av information till fler personer, men kunskapen om de behov, hinder och möjligheter som text-to-speech faktiskt innebär var dock låg när vi startade. Vi upplever dock att kunskapen ökat p.g.a. projektet. Våra presentationer på olika Wikimedia-evenemang samt *Wikimedia Diversity Conference 2017* var viktiga för att åstadkomma detta. Den sistnämnda var en stor möjlighet att få upp mångfaldsfrågor på Wikimedia-rörelsens agenda då förgrundsgestalter inom området från ett 60-tal länder samlades i Stockholm. Avsaknaden av dessa lösningar är ett hinder för informationsinhämtning för många människor. Från Wikimedia Foundations sida har det föreslagits att en långsiktig finansieringsmodell för fortsatt utveckling av Wikispeech kan vara aktuell. Detta vore en mycket positiv utveckling och vi kommer att fortsätta diskutera detta tillsammans med dem.

Kommunikationen innebär även att vi fick volontärer att intressera sig för arbetet och bl.a. en volontär från Kamerun engagerade sig i den tekniska utvecklingen och fokuserade på verktyg för att göra det enklare för volontärer att bidra till lexikonet.

Kod- och säkerhetsgranskningen visade sig dock mycket mer svår genomförd än vad vi trott. Initialt berodde det på en mycket turbulent tid hos Wikimedia Foundation där *all* den personal vi varit i kontakt med slutade. Då det saknades genomarbetade rutiner för hur ny funktionalitet som utvecklats av andra än Wikimedia Foundation kan kod- och säkerhetsgranskas var det mycket svårt att få in arbetet som en del av deras åtaganden. Detta sker för närvarande *ad-hoc* men Wikimedia Foundation arbetar aktivt med att ta fram bättre arbetsätt. Med vår iterativa process kunde vi dock fortsätta utvecklingsarbetet under hela projektet så att det finns en fungerande talsyntes för MediaWiki, även om slutgiltig kod- och säkerhetsgranskning återstår.

Vi såg även kommunikationsarbetet som ett utmärkt tillfälle för att öka medvetenheten hos den breda allmänheten om insatser rörande tillgänglighet online genom exempelvis pressmeddelanden och arbete på sociala medier.

Kommunikation gentemot forskarsamhället på olika evenemang under projektets gång var viktigt för att säkerställa att nya resurser utvecklas i framtiden och övertyga forskargrupper om att släppa material under fri licens på kort och lång sikt. Bl.a. ledde detta arbete till att vi kunde inkludera material till Wikispeech på arabiska.

8.2 Projektets budget

Redovisa projektets totala kostnad och hur det har finansierats (belopp uppdelat på finansörer) och eventuella avvikelser gentemot budget.

Under flera etapper underspenderade projektdeltagarna och pengar flyttades framåt. Då vi ursprungligen tänkt oss ett väldigt framtungt projekt innebär detta egentligen bara att arbetsbördan fördelades jämnare under projekttiden. Vi förlängde även projekttiden och sköt till egenfinansiering för att möjliggöra en del ytterligare aktiviteter.

	Budgeterat från PTS	Total kostnad	Egenfinansiering eller underspenderat
Arbetstid			
WMSE	1 120 020	1 509 300	389 280
STTS	1 146 500	1 146 500	0
KTH	391 500	1 791 500	1 400 000
Fasta kostnader			
WMSE	139 959	113 676	-26 283
TOTALT	2 797 979	4 560 976	1 762 997

Nedbruten budget per projektpart. Samtliga siffror har avrundats till hela kronor.

Wikimedia Sverige har skjutit till ca 350 000 kronor i egenfinansiering under projektets gång vilket gått till ökade kostnader för arbetstid. De minskade fasta kostnaderna är ett resultat av att vi kunde genomföra en större konferens, Wikimedia Diversity Conference, i Sverige. Detta resulterade i minskade resekostnader men medförde istället ökade kostnader för arbetstid.

KTH har skjutit till 1,4 miljoner som egenfinansiering i projektet: via en 2 års postdoc-tjänst från SRA ICT TNG har vi lagt till utveckling av metoder för att styra betoning i syntesen. Målet är att göra förståelsen av den text som läses lättare att ta till sig. Genom att göra ett system som först automatiskt genererar taggar med betoningsgraden på varje ord i en mening, som skickas med in i det neurala nätverket som genererar ljudet, så kommer vi kunna låta volontärer rätta den automatiskt genererade betoningen i en text. Dessa ändringar kan vi senare använda för att förbättra den automatiska betoningsgenereringen. I samband med detta har KTH dessutom lagt till studier där blickrörelser vid läsning av Wikipediasidor används för att skapa modeller av hur mycket fokus en läsare lägger på ord - något som samvarierar med prosodi. Finansieringen kommer från ett VR-finansierat blickforskningsprojekt, och resultaten ska publiceras under 2018.

Totalt uppgick PTS finansiering till 2 797 979 kr, den totala insatsen av projektdeltagarna inom ramen för projektet uppgår dock till 4 560 976 kr.

8.3 Projektets arbetssätt

Redogör för projektorganisation, samarbetsformer och förankring i eventuella partnerskap och intresseorganisationer.

Projektet leddes av Wikimedia Sverige som även utvecklade spelaren och hade huvudansvaret för kommunikationen. STTS ansvarade för utvecklingen av

bakomliggande funktionalitet, såsom wrappern och lexikonet medan KTH fokuserade på testning av talsyntesen samt på utveckling av prosodin.

Inledningsvis hade vi möten på veckobasis. När vi etablerat uppdelningen av uppgifter och hade en god förståelse för hur vi arbetade tillsammans började vi ha möten varannan vecka. Hela tiden höll vi task management-systemet Phabricator uppdaterat så att det gick att följa framstegen. På månadsbasis sammanställde vi en enklare rapport internt med vad som fungerat och vad som ändrats från den ursprungliga planen så att justeringar kunde ske. Baserat på detta uppdaterade vi vårt Scorecard och dessa interna månadsrapporter låg även som grund för våra Etapprapporter till PTS.

Wikimedia Sverige genomförde även möten med intresseorganisationer för personer med funktionsnedsättning för att stämma av och ta del av deras expertis om vilka funktioner i talsyntesen som var mest centrala. Baserat på den feedback vi fick från dem kunde vi omprioritera den funktionalitet vi utvecklade.

Kontinuerligt under hela projektet bedrevs ett mycket aktivt kommunikationsarbete för att säkerställa en förståelse för frågan och behovet av projektet hos både den breda allmänheten och forskare, men framförallt hos Wikimedia-rörelsen. I slutändan är det volontärerna på Wikipedia som väljer om Wikispeech aktiveras eller inte. Innan dess behöver dock koden driftsättas i Wikimedia Foundations produktionsmiljö vilket kräver en aktiv insats från Wikimedia Foundation samt uppföljning för vår del av eventuella uppkomna frågor. Den tredje orsaken till detta intensiva kommunikationsarbete var att Wikimedia-föreningar över hela världen kommer att behöva engageras för att identifiera språkresurser i deras länder. Arbetet skedde transparent så att intresserade kunde följa arbetet och engagera sig i olika insatser.

Vi deltog på flera evenemang och presenterade projektet och den internationella konferensen *Wikimedia Diversity Conference 2017* anordnades för att befästa långsiktiga strategier och investeringar i mångfald på Wikimedia-projekten. Ett stort antal av Wikimedia-rörelsens ledare befann sig på plats och Wikispeech framhölls där som ett centralt initiativ för att öka deltagandet från nya grupper.

Därtill ville vi säkerställa att kunskapen hos talsyntesforskare var hög för att förhoppningsvis kunna engagera dem i arbetet över tid och i framtida insatser. För detta har vi arbetat med att publicera fler artiklar om Wikispeech. Som ett direkt resultat av Wikispeech har en språkresurs för arabiska (en hts-röst) släppts under en fri licens.

9 Vad händer nu?

Redogör för vad som händer nu med resultatet, hur tas det om hand, startas det ett nytt projekt eller blir det en lansering. Beskriv också hur ni har spridit projektets resultat. Redogör även hur resultatet kan användas av andra som berörs av projektet även efter att projektet har avslutats. Vilka andra kan ha intresse av att ta del av projektresultatet. Dela avslutningsvis med er av tankar och visioner för framtida forskning, utveckling etc.

Mjukvaran finns tillgänglig både via Wikimedia Foundations kodrepo. Projektresultatet, talsyntesen, finns nu inkluderat på Wikimedia Foundations testserverar och driften bekostas av Wikimedia Foundation. När det senare implementeras uppbärs driftkostnaderna av Wikimedia Foundation.

Vi kommer att fortsätta utveckla Wikispeech under 2018 baserat på den feedback som ges av Wikimedia Foundations utvecklare med förhoppning att kunna inkludera Wikispeech på Wikipedia. När Wikispeech aktiveras på Wikipedia kommer projektgruppen att arbeta med en aktiv kommunikation om lanseringen genom olika kanaler. Vi genomför gärna detta i nära samarbete med PTS.

Allt producerat material och programvara är fritt licensierat och kan återanvändas av alla intresserade.

Vi identifierade flera intressanta möjligheter för fortsatt teknisk utveckling:

- Genom det nuvarande projektet har vi tagit fram en fungerande lösning, men det innebär inte att gemenskapen kommer att välja att aktivera den – projektets värde ökar extremt mycket om vi kommer till en full aktivering på Wikipedia. Vi vill därför säkerställa att vi kan hantera inkomna buggar och förbättringsförslag som kommer in under Beta så att Wikispeech kan aktiveras på Wikipedia fullt ut. Detta kommer att vara en blandning av upplevda problem och specifika önskemål som gemenskapen ser som centrala och som vi inte tänkt på.
- Vi vill bl.a. bygga vidare på rättningsfunktionaliteten så att inspelningar för att utveckla en tal-till-text-lösning kan inkluderas.
- Vi vill även se över möjligheten att skapa en metodik samt den tekniska funktionalitet som krävs för att snabbt förbättra lexikonet i samband med en kris (där information snabbt behöver nå alla i samhället på ett lättförståeligt sätt).

Utöver det rent tekniska utvecklingsarbetet ser vi ett stort behov av att utveckla metoder och material för att engagera allmänheten i att utveckla lexikonet.

Vi har sett ett stort intresse av resultatet och av kommande insatser runt insamlande av inspelningar. Vi har efter projektslutet bl.a. kontaktats av *Unicode Consortium*⁷ som är intresserade av att ta del av de fritt licensierade lexikala resurser vi skapar. Flera forskargrupper har också uttryckt intresse och även företag såsom Google. Ett fortsatt utvecklingsarbete för att säkerställa ett stort inflöde av fritt licensierat material är därför något vi ser som högprioriterat då det inte bara skulle gagna Wikipedia utan även talsyntesutvecklingen på bred front.

Gällande fortsatt forskning kommer KTH att fortsätta sin utveckling och utvärdering av datadriven syntes där man har explicit kontroll över prosodin. De två mest lovande utvärderingsmetoder som togs fram i projektet ska kalibreras

⁷ <http://unicode.org/>

för olika typer av text och behov, göras robusta, och göras fritt tillgängliga. KTH söker för närvarande medel för detta.

KTH har även för avsikt att ta fram en fritt tillgänglig svensk syntes inom ramen för Språkbanken Tal. Arbete och resurser från det tänkta fortsättningsprojektet kommer att underlätta detta betydligt.

Utöver teknisk utveckling och forskning vill vi ordna aktiviteter och andra riktade insatser för att t.ex. märka upp artiklar med Wikispeech-uppmärkning eller lägga till "Alt text" på bilder (så att det i text beskrivs vad som syns på bilden). Även skapande av utbildningsmaterial för att göra det hela skalbart är en central del. För detta skulle vi vilja sätta samman utbildningsmaterial (både instruktionssidor, instruktionsvideor, tryckt material m.m.) för att minska tröskeln för att bidra till Wikispeech och kunna aktivera fler. Därefter genomför och deltar vi på olika evenemang som bl.a. fokuserar på att skapa välfungerande exempel.

Vi vill även ge hjälp till andra organisationer som har MediaWiki att aktivera Wikispeech. MediaWiki används av tusentals olika organisationer i Sverige och utomlands och vi behöver informera dem om tillägget på olika sätt och hjälpa dem med det praktiska runt aktiveringen.