

Etapprapport

- för PTS innovationstävling

Projekttitel:	Wikispeech
Projektorganisation :	Wikimedia Sverige, med stöd av STTS AB och KTH.
Etapprapport för etapp nr:	2
Författare av detta dokument:	John Andersson
Datum för skrivande av etapprapport:	2 december 2016 (deadline framskjuten från 31 oktober 2016)
Version av detta dokument:	1
Bilaga:	Bilaga 1: Uppdaterad projektplan

1. Instruktioner

Detta dokument fylls i vid rapportering av respektive etapp och avser den aktuella etappen.

OBS! Texter skrivna kursivt ska tas bort från dokumentet och exempeltexter ska ersättas med era egna uppgifter.

2. Etappprogress

I detta kapitel beskrivs hur den redovisade etappen har gått och eventuella händelser/avvikelser samt tillhörande anledningar/orsaker.

2.1 Aktiviteter, utfall och avvikelser

Nedan beskrivs aktiviteter och utfall från etappen.

Etapp 2: Utveckling och förberedelse av utvärdering	Tidsperiod	Avvikelse i tid
Etappens tidsperiod (startdatum - slutdatum):	2016-06-01 - 2016-09-30	

Aktiviteter under etappen:	Utfall i % (vad har ni utfört av det som var planerat för denna etapp)	Avvikelse (större ändringar som gjorts - avvikelser från plan och anledning)
<ul style="list-style-type: none"> Projektledning: Etapprapport 2 inskickad (WMSE) 	Etapprapport 2 är 80 % klar	En slutgiltig uppdatering av del 2.3. <i>Kostnader för etappen</i> samt 3. <i>Uppdaterad projektplan</i> behöver ske i januari (efter överenskommelse med PTS)
<ul style="list-style-type: none"> Community konsultation: Uppsattsämnen definierade (WMSE) 	100 %	
<ul style="list-style-type: none"> Utveckling läsning: Gränsnittet för uppspelning stabilt (WMSE) 	50 %	Arbetet med uppspelningen har tagit mer resurser i anspråk än vad vi trodde. Se nedan för längre förklaring.
<ul style="list-style-type: none"> Utveckling läsning: Wrappern uppe (STTS) 	100 %	
<ul style="list-style-type: none"> Utveckling läsning: Prosodiska taggar och prosodisk styrning samt utländska ord och fraser fungerar (KTH) 	80%	Arbetet relaterat till prosodiska taggar och styrning är klart så långt

		vi planerade enligt budget i ansökan. Vi kommer slutföra crowdsourcing av prosodiska taggar på svenska samt engelska inläsningar under december då KTH har en kurs med 60 studenter som har detta som en uppgift. Utredning av behov av förbättring av utländska ord skjuts till efter att vi gjort användartester i vår på den första versionen av det svenska TTS-systemet.
<ul style="list-style-type: none"> • Utveckling förbättring: Gränsnittet för rättning stabilt (WMSE) 	25 %	Två möjliga planer finns framtagna, slutgiltigt val samt majoriteten av utvecklingsarbetet senareläggs till efter att vi deltagit på event i januari 2017.
<ul style="list-style-type: none"> • Utveckling nytt språk (WMSE) 	0 %	Då flera viktiga delar som avgör hur dokumentationen ser ut ännu inte fastslagits har vi valt att lägga de få timmarna under denna del under Etapp 3.
<ul style="list-style-type: none"> • Utveckling engelska: Språk aktiverat på Wikipedias testplattform (WMSE) 	0 %	Denna del kan inte genomföras innan Utveckling läsning är färdigutvecklat.
<ul style="list-style-type: none"> • Utveckling engelska: Språk aktiverat på Wikipedias testplattform (STTS) 	0 %	Denna del kan inte genomföras innan Utveckling läsning är färdigutvecklat.
<ul style="list-style-type: none"> • Användartester: Delrapport användartester (KTH) 	100%	
<ul style="list-style-type: none"> • Expertdiskussioner (STTS) 	100 %	
<ul style="list-style-type: none"> • Expertdiskussioner (KTH) 	100 %	
<ul style="list-style-type: none"> • Avstämningsmöte med internationella Wikimediarelsen: Reserapport (WMSE) 	100 %	

Kvarvarande aktiviteter från föregående etapper:	Utfall i % (vad har ni utfört av det som var planerat för denna etapp)	Avvikelser (större ändringar som gjorts – avvikelser från plan och anledning)
<ul style="list-style-type: none"> • • 		

2.2 Avvikelser från plan

I detta kapitel beskrivs mer detaljerat om de avvikelser (orsaker och påföljder) som gjorts från plan och som kortfattat beskrivits i föregående kapitel (kapitel 2.1).

Vi har valt att nu i de första stegen investera mer resurser i en stabilare grund för uppspelningen, för att minska de identifierade riskerna 1, 4 samt 6 (se avsnitt 3.2 nedan) samt för att snabba upp senare utveckling. Att bygga den stabilare grunden innebar dock att vi ännu inte hunnit en del av de delar vi räknat med i vår ursprungliga plan. Därtill har en del lösningar vi tittat på under förstudien visat sig olämpliga att implementera. Detta utvecklas i detalj nedan.

Detta innebär att mer av Wikimedia Sveriges utveckling sker under Etapp 3 samt 4. Vi har en del av budgeten som är oanvänd och flyttas framåt och vi planerar även att flytta över en del resurser från kommunikationsarbetet till mer utvecklingstid, då vi redan fått mycket synlighet i projektets inledning. **Vår bedömning är att vi kommer att leverera en fungerande talsyntes enligt plan, men att vi eventuellt får prioritera hårdare bland möjliga finesser på de olika komponenterna.**

För att samtliga team skall kunna bidra med sin expertis på ett bra sätt kommer vi att inleda med att aktivera för lyssning i Beta och först senare aktivera rättningar i Beta, istället för att aktivera båda delarna samtidigt. Detta möjliggör för STTS och KTH att fortsätta med majoriteten av sitt arbete och endast en del av STTS:s arbetstid och budget behöver flyttas fram till Etapp 4.

2.2.1 Revideringar gällande Utveckling läsning: Gränsnittet för uppspelning stabilt

Det tog mycket längre tid att sätta upp grunden för en tilläggsmodul än vad vi först räknat med, då det krävdes att vi satte oss in i en *ny* rekommendation (best-practise) för MediaWikis arkitektur. Det finns fortfarande en del oklarheter runt detta som tar tid att reda ut på grund av bristande dokumentation. Vi kommer att försöka bidra upstream med att rapportera de konstigheter vi

upptäckt på vägen, vilket förhoppningsvis kommer att förenkla för andra utvecklare som vill engagera sig i arbetet.

För att bygga tilläggsmodulen på ett stabilt sätt och undvika efterföljande problem valde vi att fokusera mycket mer tid och energi på att koppla in ramverken för automatiserad testning än vad vi ursprungligen tänkt oss. Vår bedömning är att det minskar risken markant att tillägget inte går igenom säkerhetsreview hos Wikimedia Foundation i ett senare skede, gör det enklare att modifiera existerande kod samt förenklar för andra att bidra med utveckling på både kort och lång sikt.

Möjligheten att återanvända existerande lösningar för att förbereda och förprocessera texten innan den skickas till talsyntesen var mycket mer begränsade än vad vi trodde i förstudien. Wikimedia Sverige var tvungna att bygga en segmenterare från grunden. Detta var ett krav för att inte behöva skicka all text på hela sidan vid varje tillfälle då det ställer orimliga krav på servrar hos WMF och dataanvändning hos användaren. Det fanns flera olika ingångar och det tog tid att experimentera och utreda vilken grund som var lämplig att utgå från (HTML eller wikitext).

I det som vi kallar gränssnitt för spelaren i vår leverabel under Etapp 2, talar vi om hela delen från frontend till backend. Vi har lagt fokus på de underliggande funktionaliteten under Etapp 1 och 2. Utöver detta har vi arbetat fram mockups för viss ytterligare funktionalitet samt för användargränssnitt. Arbetet med att implementera användargränssnitt återstår dock (d.v.s. bygga det synliga samt koppla det till bakomliggande funktionalitet). Då backend är oberoende av användarens utrustning, men frontend kan skilja sig, har fokus samt utökade resurser lagts på att så mycket som möjligt ska behandlas i backend. Det kommer att underlätta aktivering på så många typer av webbläsare och devices som möjligt.

Förprocesseringen har gjorts på ett sådant sätt att det underlättar utveckling av ljudnotifikationer som gör användaren uppmärksam på exempelvis länkar, men funktionaliteten har ännu inte färdigställts. Vi har undersökt de två möjligheter som finns för variering av uppspelningshastighet och båda fungerar med det som idag är färdigställt. Slutlig implementeringen återstår dock ännu.

Fokuset har legat på att möjliggöra uppspelning av artikeltexten då den mest centrala informationen återfinns där. Informationen på den omgivande sidan (menyer, knappar, sidfot m.m.) återstår dock. Vi uppskattar att denna del kommer att ta längre tid att genomföra då vi upptäckte att vi var tvungna att bygga Wikispeechs uppspelningsfunktionalitet på ett mer avancerat sätt än vad vi utgick från i förstudien. Dessa val gjordes för att bättre kunna bädda in den i MediaWiki.

Den spelare vi hittat under förstudien visade sig vara olämplig att använda. Genom att utveckla en egen lösning på det sätt vi gjort minskar kraven på slutanvändarens apparat och drar istället nytta av Wikimedias servrar och cachning. Det gör dock att vissa delar tagit längre tid än förväntat att bygga då vi

varit tvungna att utveckla funktionaliteten själva, däribland möjligheten att med en markering visa det ord som uttalas.

Vi har satt upp en struktur för hur vi kan utveckla möjligheten att ladda ned en ljudfil, och det finns flera närliggande lösningar implementerat i MediaWiki. Det återstår dock att bygga funktionen.

2.2.2 Revideringar gällande *Utveckling förbättring: Gränsnittet för rättning stabilt*

Vi har börjat titta på hur rättningarna kan gå till. Huvuddelen av utvecklingsarbetet återstår dock då den naturliga lösningen är att koppla den till den färdiga spelaren.

STTS har satt samman en prototyp för att lyssna på en mening och kunna bidra med rättningar. Denna lösning kan implementeras som ett verktyg för att samla in rättningar. Problemet i dagsläget är hur MediaWiki hanterar anoteringar av wikitext utan att spara dessa i texten. WMSE planerar att delta på ett utvecklar möte i San Fransisco i januari för att arbeta på en lösning med andra utvecklarteam som har samma behov. Att flera andra grupper som utvecklar MediaWiki också efterfrågar denna funktionalitet gör att vi har goda förhoppningar att detta kommer att lösas. Om detta inte går att lösa inom projektets tidsramar kommer vi att fokusera på en lösning där uppmärkning av uttal m.m. sker i wikitexten. Gemenskapen på respektive Wikipedia-språkversion kommer då att behöva diskutera och besluta vilken typ av uppmärkning som är lämplig på just deras språk vilket riskerar att förlänga tiden till full implementering efter att detta projekt avslutats.

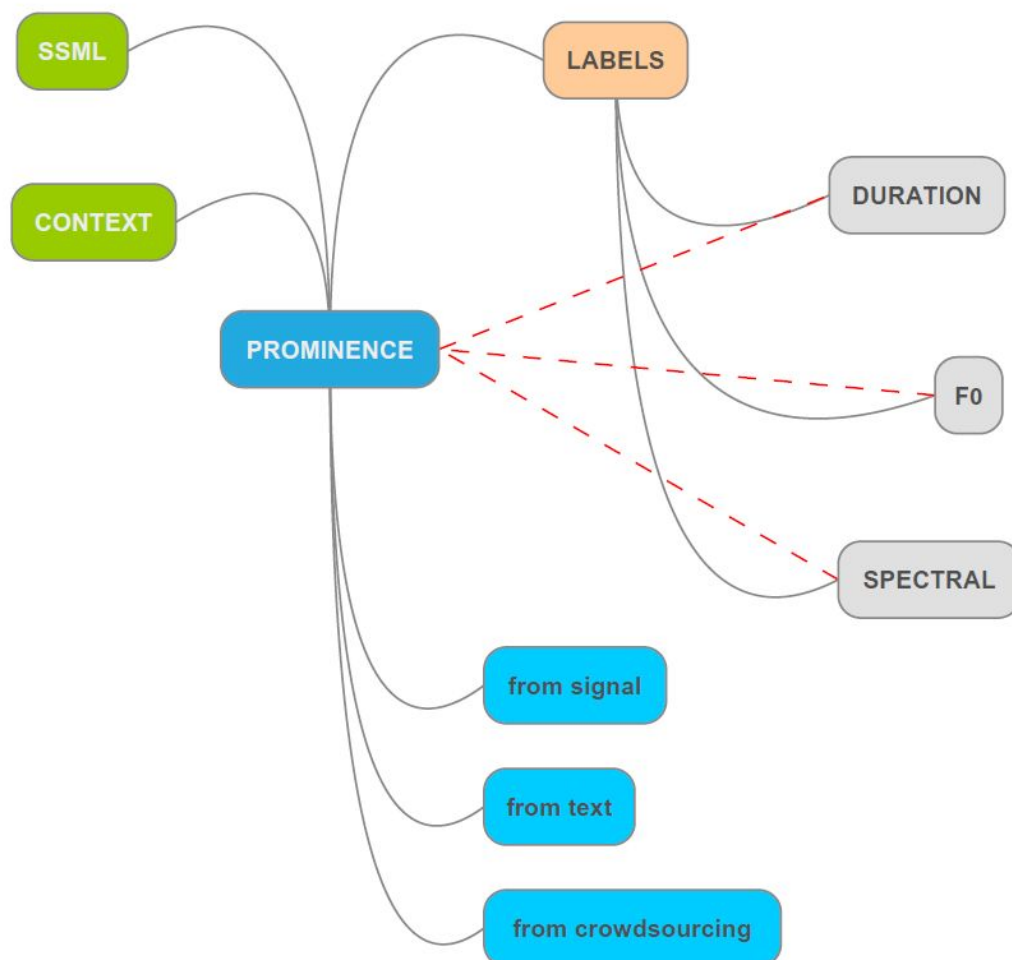
Inför mötet i januari kommer vi att titta på hur mekanismer för hur Wikipseechs behovsprofil ser ut för anoteringarna. Efter mötet i januari kommer vi att titta på hur inspelningar för att utveckla lexikonet ska gå till. Vi kommer även att implementera lagring av uppmärkningar/rättelser i artiklar på det sätt som utvecklar mötet kommer fram till.

Utvecklingen av antivandalism-stöd vid rättningar av lexikon eller artiklarna kommer att ske därefter.

2.2.3 Revideringar gällande *Utveckling prosodi*

Arbetet kring att i framtiden möjliggöra en bättre explicit kontroll över talsyntesens prosodi, d.v.s. hur systemet läser upp meningar, (betoning och fraserings), är fokuserat på två huvudegenskaper (*Möjliggöra prosodisk styrning* och *Generera prosodiska taggar*). Givet att explicit styrning av prosodi i TTS är ett olöst problem, både inom forskarvärlden och inom kommersiella system, kunde utveckling av dessa komponenter ej få plats i budgeten för detta projekt. Vad vi planerar att göra är istället att ta fram manuellt taggad data för betoning samt utreda vilka automatiska metoder som skulle kunna användas för att utveckla

prosodimoduler. KTH har även i samarbete med STTS tagit fram en tentativ TTS-arkitektur som skulle möjliggöra explicit prosodisk styrning.



Eftersom KTH har anställt en postdoc på egen finansiering (heltid under hela 18-månadersperioden) har vi dessutom kunnat lägga tid på att experimentera med hur de metoder vår utredning identifierat faktiskt fungerar på våra emfastagging av inläsningar. Hon har t.ex. testat AuToBI på preliminära prominenstagningar på svenska - det ser lovande då den faktiskt kan generera emfastaggar som överrensstämmer väl med vår manuellt taggade data.

Så snart vi har slutfört våra crowdsourcingsexperiment under december kommer hon kunna utöka testerna till engelska och större träningsmängd på svenska. Hon kommer då även kunna börja experimentera på att generera prosoditaggar från texten i dessa inläsningar och emfastaggingarna.

Under Etapp 3 kommer KTH förse STTS med träningsmaterialet från de emfastaggingar som gjordes under Etapp 2 så att de kan träna om HMM-syntesen på engelska och svenska. Detta för att möjliggöra prosodisk styrning enligt den föreslagna TTS-arkitekturen ovan. Detta ger oss möjlighet att experimentera med olika prosodiska realiseringar i de användartester vi planerar göra. Detta kommer ge oss kunskap om vikten av bra prosodi för att

målgruppen för Wikispeech-projektet skall kunna ta till sig informationen på Wikipedia-sidor.

Om det visar sig nödvändigt kommer vi i ett framtida projekt ta fram metoder för att se till att Wikipedia-sidor förses med nödvändiga taggar. Dessa kommer förmodligen initialt behövas sättas dit av Wikipedias användare så att vi kan träna en modul som sätter dit dem automatiskt.

Det vi kan få från våra emfasuppmärkningar vi gjorde under Period 2 är data på var läsaren av de texter som ingick i träningsmaterialet valde att sätta fokus. Det är inte det enda stället man kan göra det. Det beror t.ex. till stor del på kontexten (tidigare meningar eller Wikipedia-sidor som just lästs) och i talsyntesinläsningar läser man meningar i slumpmässig ordning.

Utanför Wikispeech-budgeten kommer KTH dessutom under Etapp 3 att utreda effekten av prosodiska features i DNN-baserad TTS.

2.3 Kostnader för etappen

Nedan beskrivs finansiering och kostnader som projektet haft under etappen.

	Belopp i kronor
1. Ej nyttjad finansiering från tidigare etapper	Wikimedia Sverige hade ca. 57 000 kr kvar i oupparbetad arbetstid STTS hade 58 000 kr kvar i oupparbetad arbetstid 24 192 kr (avstämningsmöte med WMF, under fasta kostnader)
2. Erhållen finansiering från PTS för aktuell etapp	611 730 kr
3. Budgeterad egen finansiering för aktuell etapp	0 kr
3. Utfall egen finansiering för aktuell etapp	0 kr
4. Utfall total kostnad för etappen	WMSE: 219 799,78 kr i oupparbetad arbetstid

	<p>STTS: 327 000 kr upparbetad arbetstid</p> <p>KTH: 147 750 kr i upparbetad arbetstid</p>
5. Fasta kostnader för etappen	23 129,25 kr i resor + annons
6. Outnyttjad finansiering vid etappens slut (som flyttas till nästa etapp)	<p>WMSE har 215 153,23 kr kvar i arbetstid till kommande etapper</p> <p>KTH har 37 500 kr kvar i oupparbetad arbetstid från Etapp 2 som flyttas till Etapp 3.</p> <p>WMSE har 29 192,75 kr kvar i outnyttjade fasta kostnader</p>

2.4 Händelser under perioden som etapprapporten avser

I detta kapitel beskrivs viktiga händelser som skett under etappen.

Fokuset under denna etapp har legat på den tekniska utvecklingen samt på förberedande arbete inför testerna.

- **Projektledning/Community konsultation/Expertdiskussioner:** Vi har fortsatt använda Phabricator som huvudsakligt verktyg för att koordinera arbetet och har varannan vecka genomfört ett sprintmöte för att diskutera hur arbetet gått samt eventuella förändringar, avvikelser eller svårigheter som uppstått. Vi har fortsatt bryta ned uppgifterna som identifierades i förstudien i mindre och mer väldefinierade uppgifter i samband med utvecklingen samt lagt in övriga projekrelaterade uppgifter där.
- **Projektledning:** Wikimedia Sverige har tagit fram licenstexter för programvaran. För säkerhets skull fick vi här stöd av Wikimedia Foundations jurister för att säkerställa bästa möjliga kompatibilitet med övrig mjukvara.

- **Projektledning:** En ny volontär har kontaktat WMSE om att engagera sig i projektet för att aktivera Wikispeech på polska och hjälpa till att utveckla lexikonet. Han har tidigare arbetat med att utveckla Samsungs talsyntes.
- **Projektledning:** Efter förfrågningar från möjliga externa partners har interna diskussioner i projektgruppen skett om hur ett eventuellt fortsättningsprojekt som fokuserar på att skapa de talresurser som behövs för nya röster samt för en open source-baserad taligenkännings-funktionalitet kan genomföras. KTH har tidigare identifierat bristen på fritt licensierade talresurser och att använda crowdsourcing kan vara en del av lösningen för att åtgärda detta. Samtal har förts med PTS om vilka möjligheter som finns runt framtida projekt. Denna möjlighet påverkar flera val av lösningar och prioriteringsordning och har därför varit viktigt att föra tidigt i projektet.
- **Projektledning:** WMSE har påbörjat en lista med Wikispeech-terminologi för att undvika möjliga missförstånd och förtydliga beskrivningarna i Phabricator.
- **Community konsultation:** WMSE har fortsatt utveckla projektrelaterade sidor på olika relevanta wikis (mediawiki.org, meta.wikimedia.org samt se.wikimedia.org). Detta för att informera aktörer om det arbete som sker.
- **Community konsultation:** WMSE har fortsatt kommunikationen med WMF:s olika team om hur vi kan samarbeta har skett fortgående, främst genom phabricator.wikimedia.org.
- **Community konsultation/Övrigt:** Möjliga uppsatsämnen kopplade till testerna har identifierats under ledning av KTH. Dessa kommer att presenterats för studenter under vårterminen.
- **Community konsultation:** Ett par olika blogginlägg om Wikispeech har skrivits för WMF:s och Wikimedia Sveriges bloggar. De har ännu inte publicerat då vi har arbetat med att ta fram en logotyp för Wikispeech.
- **Community konsultation:** WMSE deltog på konferensen Wikimania för att informera om projektet. En poster fanns uppsatt och vi deltog på både hackathon och flera workshops och presentationer som var mycket relevanta för projektet.
https://se.wikimedia.org/wiki/Projekt:Wikispeech/Global_Metrics#Wikimania_2016
- **Community konsultation:** KTH har presenterat en artikel om Wikispeech på *The 9th ISCA Workshop on Speech Synthesis* (<http://festvox.org/ssw9>). Konferensen genomförs vart tredje år och som är central för de som arbetar med talsyntes och är en viktig del för att skapa intresse från forskarsamhället som utvecklat open source syntes-plattformarna. Till konferensen togs även en poster fram som sammanfattar projektet (<https://www.dropbox.com/s/okw5w8r0rpvo4dq/poster.pdf?dl=0>). Under konferensen diskuterades projektets förlängning med ett stort antal

representanter för talsyntesutveckling, varav samtliga var positiva och de allra flesta erbjöd sig att delta i samarbeten längre fram.

- **Community konsultation:** KTH har presenterat Wikispeech-arbetet på Svensk dialogverkstad, workshop in Göteborg. Verkstaden, som är en årlig tilldragelse sedan 2006, fokuserar på samarbete mellan forskning och industri, och på praktiskt arbete. Projektet möttes med stort intresse, även från industrisidan.
- **Utveckling:** WMSE har satt upp strukturen för tilläggsmodulen baserat på de nya rekommendationerna för MediaWiki.
- **Utveckling:** WMSE har implementerat ett ramverk för automatiserad testning för tilläggsmodulen.
- **Utveckling uppläsning (navigering):** WMSE har byggt en segmenterare från grunden så att det är möjligt att skicka delar av texten till servern istället för all text.
- **Utveckling uppläsning (navigering):** Start och stopp med mus eller snabbtangenter fungerar nu och vi har arbetat mycket med att skapa möjligheten att läsa upp markerad text. Att hoppa fram och tillbaka i texten är i stort sett färdigutvecklat.
- **Utveckling uppläsning (ljudspelare):** WMSE har utvecklat en egen lösning för en spelare.
- **Utveckling uppläsning (ljudspelare):** WMSE har utvecklat den underliggande funktionaliteten för spelaren samt arbetat fram en mockup.
- **Utveckling uppläsning (ljudspelare):** WMSE har utvecklat en grund för att ljudnotifikationer ska fungera för att göra lyssnaren uppmärksam på ex. länkar eller fotnoter.
- **Utveckling uppläsning (ljudspelare):** WMSE har satt upp en struktur för hur vi kan utveckla möjligheten att ladda ned en ljudfil.
- **Utveckling uppläsning (syntesmotor):** KTH har fortsatt arbetet med att möjliggöra prosodisk styrning (emfas) och generering av prosodiska taggar:
 - manuell emfastagning (experter och/eller crowdsourcing)
 - utvärdering av två system för automatisk prosoditagning (inklusive emfas)
 - systemarkitekturdesign för möjliggörande av prosodisk styrning i TTS
 - rapportskrivande om prosodistyrning i Wikispeech
- **Utveckling uppläsning (syntesmotor):** STTS har arbetat med gränssnitt för uppdatering av uppmärkning/lexikon-syntes. Målet är en förenklad version av

Wikispeech-flödet, där en användare skall kunna ändra i lexikonet och lyssna på resultatet. Detta är en del av arbetet med att integrera de olika komponenterna. Detta har till exempel lett till ny konvertering mellan Wikispeech och MaryTTS, samt att ett par misstag i lexikonservern har upptäckts och rättats.

- **Utveckling uppläsning (syntesmotor):** STTS har utvecklat en första version av lyssna-redigera-uppdatera-loopen, implementerad så att det går att köra i en webbläsare. Vad detta betyder, är att man kan ta en text och lyssna på den i syntesen, och därpå förbättra den fonetiska uppmärkningen, och lyssna igen på resultatet. Dels kan man märka upp text med uttal direkt i texten (som SSML) och lyssna på hur ändringarna låter i syntesen, dels kan man redigera de fonetiska transkriptionerna i lexikondatabasen och spara dessa. När man sedan syntetiserar samma text igen, används de nya transkriptionerna från lexikonet. Även om detta är en ganska enkel implementation, är det intressant eftersom den i princip innehåller de steg som är grunden för hela Wikispeech-projektet (d.v.s. att en användare skall kunna redigera den fonetiska uppmärkningen, och lyssna på resultatet).
- **Utveckling uppläsning (NLP):** STTS har arbetat på automatisk validering av fonetiska transkriptioner och ”mappning” av transkriptioner (översättning fram och tillbaka mellan olika fonemuppsättningar). En del av arbetet har gått ut på att få olika komponenter att hänga ihop (exempelvis att kunna koppla en viss uppsättning valideringsregler till ett visst lexikon).
- **Utveckling uppläsning (NLP):** STTS har börjat titta på HTTPS och inloggning till lexikonservern, så att den i framtiden skall kunna göras tillgänglig via internet. (Det finns dock inte implementerat ännu.)
- **Utveckling engelska (saknade eller dåliga)/Utveckling svenska (saknade eller dåliga):** STTS har arbetat med att göra tre lexikonfiler (svenska, norska, engelska) tillgängliga i ett format som kan importeras till lexikonservern. Det går även att exportera ett lexikon från databasen till detta format.
- **Utveckling engelska (API-anpassningar/identifierar komponenter som finns)/Utveckling svenska (API-anpassningar/identifierar komponenter som finns):** HTTP-servern som anropar lexikondatabasen har strukturerats om, och några nya komponenter har tillkommit eller förbättrats (”lexicon” för uppslagning och uppdatering av lexikon; ”mapper” för att översätta mellan olika fonetiska transkriptioner; ”symbolset” som håller reda på vilka symboluppsättningar som används i olika lexikon; ”validation” för att validera ett lexikonuppslag).
- **Utveckling förbättring (rätta lexikon):** STTS har arbetat med lexikonfilformat för import av externt lexikon.
- **Utveckling förbättring (rätta lexikon):** STTS har satt ihop en rudimentär implementation av statistik för ett lexikon.
- **Utveckling arabiska:** Nawar Halabis HTS-röst för arabiska har släpps under en lämplig licens vilket möjliggör integrering i Wikispeech. Detta är ett gott

exempel på den positiva påverkansmöjlighet som finns genom denna typ av större utvecklingsinitiativ av öppen programvara.

- **Dokumentation:** WMSE har skapat en tabell med information om variabler som lokalt går att ställas in. Den är baserad på ett annat tillägg och innehåller variabelnamn, standardvärden samt beskrivningar av variablerna. Den kommer att uppdateras när nya variabler tillkommer och finns på: https://www.mediawiki.org/wiki/Extension:Wikispeech#Complete_list_of_configuration_options.
- **Dokumentation:** STTS har påbörjat beskrivning av “mock-upen” av syntesservern: https://github.com/stts-se/wikispeech_mockup/wiki
https://github.com/stts-se/wikispeech_mockup/wiki/api
- **Dokumentation:** STTS har påbörjat beskrivningen av HTTP-API:et för uttalslexikonet: <https://github.com/stts-se/pronlex/wiki/HTTP-server>

- **Expertdiskussioner:** WMSE och STTS hade ett fokuserat möte för att demonstrera delar av våra respektive system.
- **Expertdiskussioner:** KTH och STTS har haft ett par möten och pratat om ”prominens”, och hur det i framtiden skall kunna användas i Wikispeech. Med prominens menas exempelvis hur vissa delar av ett yttrande kan göras mer framträdande (exempelvis genom starkare betoning). Detta kan användas för att exempelvis hjälpa till att särskilja mellan ny och gammal information i en text. (Detta är dock inte någon trivial uppgift.)

- **Övrigt (tester):** KTH har påbörjat arbetet med testspecifikationen och bl.a. testat syntesutvärderingar på plattformen Crowdfunder. Under hösten KTH arbetat med utvecklingen av nya metoder för utvärdering av uppläsning av lång, informationsrik text. Dessa kommer att testas under 2017, bl.a. i samarbete med Myndigheten för tillgängliga medier (MTM).
- **Övrigt:** KTH har även inlett kontakter med Karolinska Institutet, Myndigheten för tillgängliga medier (MTM) samt Linköpings universitet om samarbete runt testerna genom att kombinera det med olika projekt. Utöver de initiala syntesutvärderingsexperiment som genomförs under 2017 tillsammans med MTM, sammanställs ett förslag på projektsamarbete med MTM runt resurser under samma tid.
- **Övrigt:** STTS har experimenterat med att lägga till en norsk röst för att testa hur det fungerar att skala även till andra språk. Det norska lexikonet har samma format som det svenska och har därför snabbt kunnat inkluderas. Fördelen är att genom att redan tidigt testa med ytterligare ett språk så ökar möjligheten att identifiera och åtgärda konstigheter i ett tidigt stadium.

3. Uppdaterad projektplan

PTS vill att projektet har uppdaterat dokumentet Projektplan (etapp-excelfilen, omvandlad till PDF om handläggare så önskar) för nästkommande etapp i denna etappredovisning och skickar med denna som en separat bilaga. Detta ska göras även om det inte skett några förändringar sedan förra rapporteringen.

Projektplanen ska dateras och versionshanteras.

Observera att det endast är kommande etapp(er) som ska justeras - alltså den etapp som ni nu rapporterar ska inte justeras även om den inte blev som ni planerat.

Den uppdaterade projektplanen kan t.ex. innehålla fler eller andra aktiviteter än vad som funnits med i ursprunglig projektplan samt ev. andra uppdateringar som behövs göras som åtgärder för att hantera eventuella förändringar eller avvikelser i denna etapp. Lägg även in nya rader och döpt dem "Kvarvarande aktiviteter från föregående etapper" om det är så att ni flyttar aktiviteter från etapper som redan är avslutade (se exempel i kapitel Aktiviteter, utfall och avvikelser).

3.1 Kommande händelser

I detta kapitel beskrivs kommande intressanta händelser av vikt för projektet. Nedan beskrivs även nya händelser som tillkommit sen ursprunglig projektplan.

I huvudsak kommer arbetet under Etapp 3 att fortsätta fokusera på utvecklingen av de komponenter som identifierades under förstudien.¹ Därtill kommer följande händelser av relevans att ske:

- **Projektledning:** Samtliga projektpartners kommer att diskutera vad som behöver testas med KTH, fortsätta diskutera testspeifikationerna samt påbörja arbetet med att utveckla utvärderingsmetoderna.
- **Projektledning:** Projektgruppen kommer att diskutera med olika intressenter som erbjuder sig att på olika sätt hjälpa till vad de kan bidra med.
- **Community konsultation:** Wikimedia Sverige har ansökt om deltagande för två utvecklare på Wikimedia Developer Summit i januari 2017 för att finna en lösning på hur talsyntesens uppmärkning kan behållas när texten på wikisidan ändras, samt knyta kontakter med både volontärer och anställda på Wikimedia Foundation.

¹ Se den uppdaterade Bilaga 4 Tidsuppskattningar:

<https://upload.wikimedia.org/wikimedia/se/d/d3/Wikispeech - Bilaga 4 Tidsuppskattningar-fin al.pdf> samt Bilaga 3 Etapper:

<https://se.wikimedia.org/wiki/Fil:Wikispeech - Bilaga 3 Etapper %28slutversion%29.pdf>

3.2 Uppdaterad riskanalys

I detta avsnitt beskrivs projektets risker vid etapprapporteringstillfället. Riskanalysen i tabellen nedan är uppdaterad med nya risker som identifierats sen förra etapprapporteringen. Även de gamla riskerna bedöms och viktas på nytt.

Låt även gamla risker som bedöms vara avklarade stå kvar i tabellen.

Version av riskanalysen (v.1 är den som var med i förstudieoriginalet):	1.1
Beskriv vad som är uppdaterat/förändrat från förra versionen:	Inga förändringar sedan Etapprapport 1.

Risk nr.	Risk	Sannolikhet (S)	Effekt (E)	Risikfaktor (S)*(E)
1	Att vår kod fastnar i code review/security review hos Wikimedia Foundation	2	2	4
2	Att de val vi gjort runt utveckling inte överensstämmer med vad slutanvändarna vill se	2	2	4
3	Att Wikimedias volontärgemenskap har åsikter om vad som bör göras vilket inte överensstämmer med vad slutanvändarna vill se	3	2	6
4	Komponenter vi identifierat under förstudien visar sig vara otillräckliga när vi implementerar dem på en av världens största webbplatser	3	2	6
5	Att talsyntesen inte är tillräckligt bra för att slutanvändarna skall vilja använda den (dåligt uttal på ord m.m.)	2	2	4
6	Att Wikimedia Foundation ej anser att mjukvaran är stabil nog för att implementeras i produktion. Detta påverkar risk 1.	2	3	6
7	Att de olika Wikipedia-gemenskaperna (en pe	2	3	6

	språk) inte önskar aktivera Wikispeech			
8	Att Wikispeech är så krävande att Wikimedia Foundations servrar och resurser ej kan bära det	1	2	2

Risk nr.	Förebyggande åtgärd (Beskrivningen hur risken hanteras)
1	<p>Redan under förstudien påbörjade vi samtalen med Wikimedia Foundation om projektet och började sätta upp strukturen på Phabricator för att de redan från dag ett av projektet skall kunna planera in det i sitt arbete.</p> <p>Vi har fortsatt prata med Wikimedia Foundations personal vilket varit viktigt då flera av de på chefsnivå i organisationen har slutat. Sedan en tid har de dock stabiliserat sig och vi bedömer därför risken som tidigare.</p>
2	<p>Vi har under förstudien initierat kontakter och tagit in deras tankar och åsikter för att minimera risken.</p> <p>Vi kommer att ha flera utvärderingstillfällen med representanter från funktionsnedsättningsorganisationerna. Då processen är öppen har alla målgrupper möjlighet att närsomhelst under utvecklingen delge sina tankar och förslag. Vi förväntar oss dock inte att de själva skall hålla koll utan vi sammanställer frågor till dem som vi förmedlar allt eftersom. Vi har redan under Etapp 1 bett om feedback på ett antal frågor vi brottas med.</p>
3	<p>För att göra större förändringar på Wikimedias projekt krävs konsensus i den gemenskap av volontärer som bidrar till dem.</p> <p>För att minska risken att de kommer fram till beslut som minskar värdet av produkten måste vi vara väldigt tydliga i vår kommunikation och se till att ha välunderbyggda slutsatser, tillgängliggöra allt vårt beslutsunderlag kontinuerligt samt lyssna och besvara kommentarer och frågor väldigt snabbt.</p>
4	<p>Vi har tagit höjd för detta i vår tidsplanering genom att använda en beprövad modell för att beräkna tidsåtgång för ett utvecklingsprojekt.</p>
5	<p>Vi har satt av relativt mycket resurser för att säkerställa att talsyntesen redan från början håller en relativt god nivå.</p> <p>Vi tror även att möjligheten att åtgärda irriterande felaktigheter på egen hand minskar risken att användarna slutar använda</p>

	Wikispeech, även om det inte är perfekt första gången de använder den.
6	Vi kommer först att aktivera Wikispeech på Wikipedias testplattform för att testa att allting fungerar. När vi väl kommit så långt sjunker sannolikheten till en 1:a.
7	Vi kommer att aktivera Wikispeech som en Beta Feature på Wikipedia. Detta innebär att inloggade användare kan pröva funktionaliteten och ge feedback innan Wikispeech aktiveras för alla.
8	Vi kommer inledningsvis att aktivera Wikispeech på Wikipedias testplattform. Detta ger oss och Wikimedia Foundation en möjlighet att bevaka vilka resurser Wikispeech kräver och tidigt åtgärda eventuella resurskrävande funktioner.