

# Etapprapport

## - för PTS innovationstävling

---

<b>Projekttitel:</b>	Talresursinsamlaren – För ett tillgängligare Wikipedia genom Wikispeech
<b>Projektorganisation:</b>	Wikimedia Sverige, med stöd av STTS AB och KTH.
<b>Etapprapport för etapp nr:</b>	4
<b>Författare av detta dokument:</b>	André Costa et al.
<b>Datum för skrivande av etapprapport:</b>	2021-03-09
<b>Version av detta dokument:</b>	1.0
<b>Bilaga:</b>	Bilaga 1: Budget och etapper

## Innehållsförteckning

<b>1. Instruktioner</b>	<b>3</b>
<b>2. Etappprogress</b>	<b>3</b>
2.1 Aktiviteter, utfall och avvikelser	3
2.2 Avvikelser från plan	5
2.3 Kostnader för etappen	6
2.4 Händelser under perioden som etapprapporten avser	7
<b>3. Uppdaterad bilaga – Budget och etapper</b>	<b>9</b>
3.1 Kommande händelser	10
3.2 Uppdaterad riskanalys	10

## 1. Instruktioner

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

## 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.

Hämta aktiviteter från senaste uppdaterade Projektplan. Redovisa progress/utfall och avvikelser i tabellen nedan. Utfallet ska redovisas i % av utförd aktivitet (OBS! inte % av mängden förbrukade timmar).

Om det finns "rester" av aktiviteter från föregående etapp/etapper så lägger ni dem i tabellen "Kvarvarande aktiviteter från föregående etapper" och redovisar på samma sätt.

Här följer ett exempel på hur detta kan se ut. **OBS! Texter skrivna kursivt ska tas bort från dokumentet och exempeltexter ska ersättas med era egna uppgifter.**

<i>Etapp 2: Utveckling och förberedelse av utvärdering</i>	<i>Tidsperiod</i>	<i>Avikelser i tid</i>
<b>Etappens tidsperiod (startdatum – slutdatum):</b>	201101-210301	

<b>Aktiviteter under etappen:</b>	<b>Utfall i % (vad har ni utfört av det som var planerat för denna etapp)</b>	<b>Avvikelser (större ändringar som gjorts – avvikelser från plan och anledning )</b>
<ul style="list-style-type: none"> <li>Projektledning: Produktbeskrivning på mediawiki.org; Etapprapport 4 inskickad</li> </ul>	Produktbeskrivning på mediawiki.org: 100 %  Etapprapport 4: 100 % klar	
<ul style="list-style-type: none"> <li>Community konsultation: Mailutskick på samtliga relevanta maillistor; wikisidor uppdaterade</li> </ul>	Mailutskick på samtliga relevanta maillistor: 75 % klar	Samtliga förberedelser gjorde men utskicket behöver anpassas för svenskspråkig målgrupp.

## Etapprapportering

	Wikisidor uppdaterade: 100 % klar	
• Kommunikation	100? % klart	
• Kommunikation: Deltagande på konferenser	25% klart	På grund av pandemin har antalet möjliga konferenser varit färre, som ett resultat av detta har mer arbetstid lagts på andra aktiviteter.
• Integrering med spelare (och Wikimediaplattformarna)	100 % klart	
• Utveckling Inspelare: Fungerande första version	75 % klart	
• Utveckling Automatisk validerare: Fungerande första version	75 % klart	Mycket finns, saknas testning och konfigurering.
• Utveckling Lagring: Fungerande lagring på plats	90 % klart	Intern lagring är i princip klar, enbart en mindre bugg återstår att åtgärda.
• Utveckling Granskare: Fungerande första version	20 % klart	Interna API:er och datamodeller är färdiga men gränssnittet har ej påbörjats.
• Utveckling Annoterare: Fungerande gränssnitt på plats	25 % klar	Underliggande lagring har förberetts men ej ett gränssnitt integrerat med Talresursinsamlaren.
• Utveckling Annoterare: Fungerande första version	75 % klar	Mycket finns, saknas testning och konfigurering.
• Utveckling Exporterare: Grundläggande funktionalitet	25 % klart	Se reviderad plan.
• Utveckling Manus: Färdig komponent och exempelmanus på plats	90 % klart	Viss dokumentation samt integration med Talresursinsamlaren kvarstår
• Användartester	100 % klart	
• Expertdiskussioner/rapportskrivande	100 % klart	
• Arbetstid för referensgrupp, evenemang samt kommunikation	0 % klart	Har p.g.a. förseningar skjutits till nästa etapp.

Kvarvarande aktiviteter från föregående etapper:	Utfall i % (vad har ni utfört av det som var	Avvikelser
--	---	------------

	planerat för denna etapp)	(större ändringar som gjorts - avvikelser från plan och anledning )
• N/A		
• N/A		

## 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).

- Integrering med spelare och Wikimediaplattformarna: Vi har tidigare avgjort att en lansering av text-till-tal funktionaliteten på (svenskspråkiga) Wikipedia är centralt avgörande för att ett stort intresse för Talresursinsamlaren ska kunna skapas hos Wikimediagemenskapen. Det är förbättringar av talsyntesen och skapandet av nya röster för denna som är det naturligaste sättet att visa på de direkta effekterna av en inspelningsinsats. Som ett resultat av detta har vi redan tidigare valt att prioritera upp arbete med denna komponent, på bekostnad av viss funktionalitet i Talresursinsamlaren. Beslutet att lansera Wikispeech (text-till-tal) på Wikipedia görs av Wikimedia Foundation och under etappen fick vi bekräftat att detta ska göras (se *Händelser...* nedan för mer information). Men lanseringen är begränsad till en pilot specifikt för svenskspråkiga Wikipedia.

Wikimedia Foundation gjorde det även tydligt att deras insats i dagsläget är begränsad till stöd under den initiala lanseringen. Vi ser därmed att den initiala lanseringen måste innehålla all funktionalitet som är central för övertyga gemenskapen om värdet med en integrerad talsyntes och därmed utöva det externa tryck på Wikimedia Foundation som krävs för att ytterligare stöd för funktionaliteten ska prioriteras i framtiden. En saknad central komponent identifierades (lexikonförbättring) och arbetet med denna prioriterades därför upp under denna och kommande etapp. Detta fick till följd att leverabler för Talresursinsamlaren har försenats ytterligare och att dessa i vissa fall behövt revideras.

Detta har haft direkta effekter på *lanseringen av spelaren*, lanseringen av en *demoinstans av Talresursinsamlarplattformen* och den *kommunikation mot gemenskapen* dessa två var kopplade till. Alla tre var händelser som i förra etapprapporten förväntades ske under denna etappen, men nu tvingats skjutas på till nästa etapp.

- Resor och evenemang: Trots att kostnaderna för resor och evenemang tidigare reviderades ner har den fortsatt allvarliga Corona-pandemin inneburit att inte ens de minskade medlen kunnat konsumeras. Vi föreslår därför att flexibilitet medges om att, vid behov, föra över delar av de medel som under Etapp 5 avsatts för rese- och evenemangskostnader till arbetstid. Vi kommer att, tillsammans med Dyslexiförbundet, undersöka hur vi kan

organisera de planerade inspelningsevenemangen, men utgår från att dessa kommer att genomföras på distans.

- Arbets tid för referensgrupp, evenemang samt kommunikation: Den fördröjning av Talresursinsamlaren som implementationen av lexikonförbättring inneburit har gjort att referensgruppens utvärdering samt insatserna med event och kommunikation har tvingats skjutas till nästa etapp.
- Manus: Manuset är i praktiken färdigt. Tyvärr innebar en del förseningar att integrationen med Taldatainsamlarplattformen samt exempelmanuset inte hann slutföras under etappen. Dessa genomförs i början av nästa etapp.
- Annoterare: Mycket är på plats. En svårighet har varit att den funktionalitet som erbjuds är starkt kopplad till de gränssnittsbegränsningar som finns. utan att helt ha kartlagt dessa är det därför svårt att testa och en del gissningar och uppskattningar blir nödvändiga. I början av nästa etapp kommer gränssnittsbegränsningarna att utredas varpå en fungerande första version kan slutföras.
- Granskaren: Interna API:er och datamodeller har skapats för granskningsprocessen men inget arbete har gjorts för gränssnitt eller publika API:er under etappen. Detta då både full integrering av manus samt fullt fungerande inspelning behöver vara på plats för att effektivt kunna testa komponenten.
- Reviderade utvecklingsmål: Som ett resultat av den tidsåtgång som aktiviteten *Integrering med spelare och Wikimediaplattformarna* tagit i anspråk har några av de planerade aktiviteterna reviderats.

*Lagring*: Denna komponent är nu begränsad till intern lagring därmed kommer inte verktyg för import av taldata samt integrerade flöden för lagring på externa plattformar att utvecklas inom ramen för projektet.

*Exporterare*: Inget dedikerat verktyg för att "Ladda ner arkiv" byggs. Istället kommer dokumentation att skapas som beskriver hur existerande mekanismer och verktyg för MediaWiki kan användas för att göra datauttag.

### 2.3 Kostnader för etappen

Nedan beskrivs finansiering och kostnader som projektet haft under etappen.

Finansiering	Belopp i kronor
Ej nyttjad finansiering från tidigare etapper (PTS samt egen finansiering)	0 kr
Erhållen finansiering från PTS för aktuell etapp	619 013 kr
Budgeterad egen finansiering för aktuell etapp	227 680 kr
<b>Total finansiering för etappen</b>	<b>846 693 kr</b>
Kostnad	Belopp i kronor

Utfall egen finansiering för aktuell etapp	WMSE: 129 996 kr STTS: 15 210 kr KTH: 82 474 kr Dyslexif.: 0 kr Total: 227 680 kr
Utfall PTS finansiering för aktuell etapp	WMSE: 387 001 kr STTS: 154 790 kr KTH: 52 701 kr Dyslexif.: 0 kr Total: 594 492 kr
<b>Utfall total kostnad för aktuell etapp</b>	<b>822 172 kr</b>
<b>Outnyttjad finansiering vid etappens slut som flyttas till nästa etapp (Total finansiering för etappen – Utfall total kostnad för aktuell etapp)</b>	<b>24 521 kr</b>

## 2.4 Händelser under perioden som etapprapporten avser

I detta kapitel beskrivs viktiga händelser som skett under etappen.

Fördjupade diskussioner med WMF om integration: Under etappen har vi haft flera diskussioner med Wikimedia Foundations *Chief Product Officer* kring integration av spelaren och talresursinsamlaren. Direktsamtal på denna nivå var tacksamt då vi tidigare fått motsägelsefulla budskap från dem kring lansering av Wikispeech, något som lett till förseningar hos oss. Under dessa samtal beskrev de hur deras strategi ändrats och att de nu aktivt försöker begränsa antalet nya tjänster som driftsätts (riskanalysen har uppdaterats som ett resultat av detta). Detta är en förändring mot tidigare där säkerhetsgranskingen identifierats som flaskhalsen. Trots detta fick vi intygat från dem att Wikispeech (text-till-tal) ska kunna driftsättas som en pilot specifikt för svenskspråkiga Wikipedia, inom tidsspannet för det här projektet.

Vi har även förberett en *Performance Review* som gör det möjligt för testarna hos Wikimedia Foundation att utvärdera den påverkan Wikispeech har på driften av Wikipedia samt förklarar de tekniska vägval som gjorts och den belastningsminimering som genomförts.

Kommunikation mot Wikimediagemenskapen: Under etappen påbörjades kommunikativa förberedelser inför en global lansering av Wikispeech. Vi förberedde även insatser för att få stöd hos den globala gemenskapen för Wikispeech som funktion. Detta innebar dels utveckling av kommunikations- och volontäringagemangsplaner, dels inledande kontakter med rätt personer inom rörelsen för att kunna skala upp arbetet med att få stöd för funktionen inom rörelsen. Detta arbete pågick mellan november och januari, innan indikationer inkom på att Wikispeech i första hand först kommer att aktiveras som pilot på svenskspråkiga Wikipedia först, varför delar av dessa planer har börjat styras om mot en i högre grad svensk målgrupp. Det arbetet kommer att fortsätta under nästa etapp.

Integrering med spelare och Wikimediaplattformarna: En modell för lexikonförbättring har tagits fram och börjat implementeras. Denna komponent, att kunna förbättra hur en artikel låter, är den saknade komponent (se *Avvikelser...* ovan) som vi avgjorde måste finnas med redan vid lanseringen. Grunden i modellen är att använda sig av de inbyggda funktionaliteterna för versionshantering som finns i MediaWiki för att där hålla ett förbättrat lokalt lexikon, förbättringarna kommuniceras sedan till det större lexikon som används av talsyntesen. Genom att hålla dessa lexikon åtskilda kan de välfungerande mekanismer för hantering av vandalism etc. som finns i MediaWiki nyttjas samtidigt som talsyntesen och dess lexikon förblir en generaliserad komponent som kan användas både av MediaWiki och annan mjukvara. Modellen har samtidigt gjorts flexibel nog att den kan integreras med t.ex. Wikidata i framtiden eller hållas helt intern för en wiki.

Parallellt med detta har vi fortsatt kodmodernisering och belastningsminimering samt förbättrat dokumentationen för slutanvändare. Ett särskilt fokus har även lagts på att eliminera eventuell negativ påverkan som talsyntesen skulle ha för en användare som inte är intresserad av att nyttja den.

Manus: Under etappen har manusverktyget utökats med mer dokumentation, demofunktionalitet samt mindre förändringar för att förenkla dess användning. Manusverktyget är i princip färdigutvecklat sånär som på viss dokumentation och slutlig integrering med Talresursinsamlaren.

Annoteraren och Automatisk validering: API:et för Annoteraren har utvecklats vidare med tidskoder på olika nivåer (stycke, mening, ord, fonem). Automatisk validering har påbörjats, med hjälp av ASR och forced alignment. Resultatet är i dagsläget experimentellt och inte redo för produktion.

Parallellt med detta arbete har funktionalitet implementerats för lagring och versionshantering av den data som genereras av Annoteraren. Lagringen har integrerats nära med MediaWiki för att kunna dra nytta av de inbyggda mekanismer som finns för hantering vandalisering. Samtidigt innebär en nära integrering att grundläggande exportfunktionalitet direkt blir tillgänglig genom de metoder som redan existerar för MediaWiki.

Talresursinsamlaren: Plattform via vilken talresursinsamlingen sker och som knyter samman de olika komponenterna har utvecklats vidare för att bli mer robust genom att återanvända eller utöka nyare funktionalitet i MediaWiki. Fokus under etappen låg på integrering av Inspelaren samt lokal lagring av det inspelade talet.

Utvärdering: Arbetet med validering av talsyntes för långa och innehållsrika texter har löpt vidare enligt plan, med mindre modifikationer som tidigare noterats p.g.a. Covid-19, som omöjliggör t.ex. tester med fokusgrupper. KTH har tagit fram tre metoder som syftar till att undersöka olika aspekter av syntes i Wikispeech-sammanhang. I första hand den ARS-baserade (Audience Response System) metod som tidigare beskrivits. Här förbereds ett omfattande valideringsexperiment som jämför metoden med resultat från the Blizzard



Challenge<sup>1</sup>, talsyntesvärldens de facto utvärderingsstandard. Blizzards metod ger inte speciellt bra information vad gäller lång och informationsrik text, men våra nya metoder behöver ändå kunna mäta sig med de gängse metoderna.

En andra metod som står klar för testning i större sammanhang bygger på att mäta den tid det tar att tillgodogöra sig information ur en text. Metodens stora styrka är att den klarar att jämför olika modaliteter och metoder för informationsinhämtning från text på ett rättvisande sätt. Inledande studier ger hoppfulla men något tvetydiga resultat, och en större studie förbereds under våren.

Slutligen har KTH tagit fram en metod för att avgöra hur pass nära en mänsklig inläsare syntes kommer. Metoden används i en första större testserie i mars 2021, och resultaten kommer att publiceras vid lämpligt tillfälle under samma år. Nyckeln i metoden är att variera mängden ljud lyssnarna för höra, och använda den mängd som behövs för att uppfatta om det är syntetiskt tal eller mänsklig inläsning som metrik för allmän talkvalitet. Metoden ger oss ett differentierat mått som ger möjlighet att bättre analysera vad som spelar roll vid syntesträning.

Dissemination: Då dissemination vid fysiska konferenser även fortsatt är olämpligt arbetar KTH nu med flera rapporter som kommer att publiceras dels på distanskonferenser och workshops under året, dels i en sammanfattande rapportform.

GDPR: Tolkningen av GDPR kring taldata är fortfarande oklar, trots analyser av jurister på flera myndigheter (bland annat KTH och MTM). KTHs Språkbanken Tal gör under 2021 en större genomlysning av lagen med hjälp av KTHs jurister samt ett omfattande underlag från en motsvarande genomlysning som gjordes av finska Språkbanken under 2020. Det är dock oklart om resultatet kommer att vara redo under nästa etapp. Parallella satsningar för att få klarhet i hur taldatainsamlingar kan göras tillförlitliga, tillgängliga och långsiktiga genomförs även av andra aktörer i Sverige.

I dagsläget är det dock inte klart hur taldata kan hanteras. En utväg som vi tittar på i projektet är att träna med generella modeller, till exempel av typen Wav2Vec2<sup>2</sup>. Dessa kan inte återföras till individ, men de kan ändå göra stor nytta vid t.ex. talsyntesframställning.

### 3. Uppdaterad bilaga – budget och etapper

PTS vill att projektet har uppdaterat dokumentet Budget och etapper för nästkommande etapp i denna etapp redovisning och skickar med denna som en separat bilaga. Detta ska göras även om det inte skett några förändringar sedan förra rapporteringen. Dokumentet ska dateras och versionshanteras.

---

<sup>1</sup> <http://festvox.org/blizzard/>

<sup>2</sup> <https://arxiv.org/abs/2006.11477>

**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 dokumentet kan t.ex. innehålla fler eller andra aktiviteter än vad som funnits med ursprungligen samt ev. andra uppdateringar som behövt göras som åtgärder för att hantera eventuella förändringar eller avvikelser i denna etapp. Lägg även in nya rader och döp dem "Kvarvarande aktiviteter från föregående etapper" om det är så att ni flyttar aktiviteter från etapper som redan är avslutade (se exempel i kapitel 2.1).

### 3.1 Kommande händelser

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

Integration av spelaren: Under den kommande etappen förväntar vi oss att spelaren kommer att finnas tillgänglig som en betafunktion på svenskspråkiga Wikipedia tillgänglig för inloggade användare.

Demoinstans av Talresursinsamlarplattformen: En publikt tillgänglig demoinstans av Talresursinsamlarplattformen kommer att ha publicerats vilken inkluderar merparten av de utvecklade komponenterna. Via denna kommer initial återkoppling att tas in.

Kommunikation mot gemenskapen: En engagemangskampanj riktad mot gemenskaperna på Wikimediaplattformarna kommer att genomföras inför beta-lanseringen av spelaren. Som en del av denna kommer vi även att skapa engagemang för Talresursinsamlaren.

Metodologi för inspelningsevent: Tillsammans med Dyslexiförbundet tar vi fram en detaljplan för genomförande och utvärdering av inspelningsevent. Då fokus är på metodologin kan tester av inspelning komma att ske med andra verktyg än bara Talresursinsamlarplattformen för att göra processen mindre känslig för ytterligare förseningar.

Hantering av användare i utvärderingar: De metoder för talsyntesutvärdering som utvecklas på KTH resulterar i webbaserade tester som i första hand drifas på KTHs Språkbanken Tal. Rekrytering sker både genom Wikimediaplattformarna och på andra sätt. En användartestplan som beskriver hur användardata hanteras under olika omständigheter ska förfärdigas för att säkerställa att användare inte kommer i kläm.

### 3.2 Uppdaterad riskanalys

I detta avsnitt beskrivs projektet 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.

<b>Version av riskanalysen (v.1 är den som var med i förstudieorginalet):</b>	v.5
<b>Beskriv vad som är uppdaterat/förändrat från förra versionen:</b>	Effekten av risk 4 har ökat för att inkludera effekterna av de utvecklingsinsatser detta krävt/kräver och den arbetstidsåtgång detta medfört.

Så här vill PTS att man beskriver riskerna:

**Sannolikhet** – hur stor är sannolikheten att denna risk kommer att slå in, 1 är låg sannolikhet och 3 är hög sannolikhet.

**Effekt** – hur stor blir effekten om den identifierade risken blir aktuell, 1 är liten effekt och 3 är stor effekt.

Risker med riskfaktor 9 bör elimineras, risker med riskfaktor större eller lika med 4 ska elimineras, minimeras eller hanteras. Övriga risker ska bevakas.

Risk nr.	Risk	Sannolikhet (S)	Effekt (E)	Riskfaktor (S)*(E)
1	GDPR och annan lagstiftning omöjliggör helt fri delning av inspelat tal	2	2	4
2	Basteknologier från Språkbanken Tal hinner inte bli färdiga i tid för att komma projektet till nytta under projekttiden	3	1	3
3	Skyddet för användares integritet hos Wikimedia orsakar brister i den demografiska metadata vilket	2	1	2

## Etapprapportering

	försämrar taldataets användbarhet			
4	Kod- och säkerhetsgranskning av nyutvecklad mjukvara sker inte och denna kan därför inte gå live	3	2	6
5	Våra teknologier tar för stora tekniska resurser i anspråk för att accepteras av Wikimedia Foundation.	1	1	1

Förebyggande åtgärder ska beskrivas utförligt (för risker  $\geq 4$ ).

Risk nr.	Förebyggande åtgärd (Beskrivningen hur risken hanteras)
1	<p>Det råder ännu osäkerhet kring tolkningen av den nya lagen GDPR, så vi måste planera in en viss flexibilitet. En del av projektgenomförandet handlar därför om att utreda om anonymisering av röst och användardata behöver ske. Om lagstiftningen förhindrar lagring och delning av ren röstdata ökar vi fokus på anonymisering och tillför resurser (som tas från inspelat tal).</p> <p>Fokus under projekttiden kommer därför inte vara den aktiva insamlingen av inspelat tal utan att bygga verktygslådan (inklusive anonymisering) för att omgående kunna påbörja insamling när lagen förtydligas. Därtill kan verktyg som fokuserar på insamlande av annoteringar etc. nyttjas utan problem då de inte samlar in uppgifter som kan identifiera en person.</p>
2	Risikfaktor under 4. Språkbanken Tals basteknologiutveckling har i det närmaste helt lagts på is under pandemin, då den bygger på inspelningar av stora mängder tal i studio, vilket

	<p>inte är möjligt i dagsläget. Risken ger dock marginella och tillfälliga effekter på projektgenomförandet.</p>
3	<p>Risikfaktor under 4. Verktyg och metoder tas fram så att de ger stöd för användandet av demografiska metadata, men fungerar även utan. Även med begränsad mängd demografisk data har det insamlade datat ett mycket stort värde för många områden.</p>
4	<p>Risken har minimerats genom att projektet förutom med Wikimedia Foundation har etablerat samarbete med Wikimedia Deutschland som har ett stort utvecklarsteam med nödvändiga erfarenheter och behörigheter. En engagemangskampanj riktad mot Wikimediagemenskapen säkerställer att det kommer finnas tryck utifrån för att Wikimedia Foundation ska genomföra sitt åtagande parallellt med att vi bearbetar beslutsfattare på olika nivåer inom organisationen.</p> <p>Effekten har minimerats ytterligare genom att Talresursinsamlaren har designats så att den ska kunna köras vid sidan av Wikimedia Foundations plattformar om behovet skulle uppstå. Effekten för integrationen av Spelaren på Wikimediaplattformarna kvarstår dock och därmed även en indirekt effekt på Talresursinsamlaren i formen av förlorad arbetstid.</p>
5	<p>Risikfaktor under 4. Risken har minimerats genom att vi tidigt under projektet tar fram uppskattningar för de resurser vi behöver och för en dialog med Wikimedia Foundation om dessa. Parallellt med detta implementerar vi den best practice som rekommenderas av Wikimedia Foundation för att minimera våra resursbehov.</p>