

Samverkansprojekt mellan
Tillväxtverket och ESV



EKONOMISTYRNINGSVERKET

Rapport

Vidareutveckling av applikationen Hitta remissvar

En tillämpning av text som data i staten och test
av AI-verktyg

Publikationen kan laddas ner
från ESV:s webbplats esv.se.

Datum: 2024-02-27

Dnr: 2023-07346-4

ESV-nr: 2024:22

Copyright: ESV

Rapportansvarig: Kristin Lilieqvist

Förord

Denna rapport har tagits fram av Kristin Lilieqvist, Pierre Mesure och Oskar Weinberger på Ekonomistyrningsverket och Robert Kron, Tora Philp och Emma Wallerö på Tillväxtverket.

Vi vill tacka alla personer från olika myndigheter som under projektets gång bidragit i våra workshoppar, användartester, enkäter och demonstrationer samt lämnat värdefulla synpunkter på prototypen eller rapporten.

Stockholm
2024-02-27

Martin Sparr

Enhetschef

Kristin Lilieqvist

Utredare

Innehåll

Sammanfattning	6
Specifikt mål men omfattande syfte	6
Många användarkontakter och ny funktionalitet	6
Möjligheter att permanenta och/eller bygga vidare på Hitta remissvar	7
Lärdomar och rekommendationer för bättre nyttjande av textdata och AI-verktyg i offentlig sektor	7
1 Inledning	8
1.1 Bakgrund	8
1.1.1 Tidigare arbete och ursprunglig prototyp	8
1.1.2 Snabb AI-utveckling i omvärlden	9
1.1.3 Kopplingen till Tillväxtverket och ESV:s uppdrag	9
1.1.4 Möjliggörande aktörer	10
1.2 Projektets mål och syfte	11
1.2.1 Genomförande och avgränsningar	11
1.2.2 Kommunikation av resultatet	12
1.3 Disposition och läsanvisning	12
2 Utvärdering, testning och användarfeedback	14
2.1 Demonstrationer	14
2.2 Enkät	14
2.3 Testning av klassificering och matchning med hjälp av AI	15
2.3.1 Klassificering, sentimentanalys och sammanfattningar med hjälp av generativ AI	16
2.3.2 Matchningar mellan förslag och remissvar	19
2.4 Förenkling av användargränssnitt	19
2.5 Summering av återkommande synpunkter	19
2.6 Faktisk användning av Hitta remissvar	20
2.6.1 Besöksstatistik	20
2.6.2 Hitta remissvar som en del av remissvarsrutinen	21
3 Prototypens nya funktionalitet	22
3.1 Grafisk identifiering av förslagsrutor för ökad exakthet	22
3.2 Göra remissvaren sökbara	23
3.2.1 Hantera risker med personuppgifter	23
3.3 Möjliggöra sökning i utredningarnas konsekvensbeskrivningar	26
3.4 Matchningar mellan förslag, remissvar och konsekvensutredningar	27
3.4.1 Matchningar inom samma datatyp	27
3.4.2 Matchningar mellan olika datatyper	27
3.5 Automatisk uppdatering av datakällorna	29
3.6 Länkar till sparade sökningar	29
3.7 Övrig vidareutveckling av användargränssnittet	29

3.8	Ytterligare funktioner för att förbättra Hitta remissvar	30
3.8.1	Koppla fler datakällor till applikationen	30
3.8.2	Tillhandahålla API	30
3.8.3	Inkludera äldre remissvar	30
3.8.4	Inbyggd funktion för användarfeedback.....	30
3.8.5	Sortera bort remissvar utan synpunkter.....	31
3.8.6	Remissammanställning, gruppering och identifiering av mönster i data	31
3.8.7	Ytterligare förfining av användargränssnittet	31
4	Att gå från prototyp till färdig tjänst	32
4.1	Ytterligare anpassningar krävs	32
4.2	Kostnader för att permanenta tjänsten.....	33
4.2.1	Utvecklingskostnader	33
4.2.2	Löpande förvaltning och driftskostnader.....	33
4.2.3	Infrastrukturkostnader.....	33
5	Varför öppna data och öppen källkod ännu inte är fullt implementerade	35
5.1	Öppna data.....	35
5.2	Öppen källkod.....	35
6	Lärdomar och rekommendationer för användning av textdata och AI-verktyg i offentlig sektor	37
6.1	Databeredskapen måste bli bättre för att ta tillvara på potentialen i offentlig text	37
6.1.1	Några av nuvarande problem	37
6.1.2	Budskap för bättre databeredskap	38
6.1.3	Åtgärder som enskilda remissinstanser kan göra för att bidra till bättre databeredskap och funktionalitet i Hitta remissvar	39
6.2	Stort omvärldsberoende och hastigt ändrade förutsättningar	40
6.3	Även ofärdiga och ickeperfekta prototyper kan ge konkret nytta	41
6.4	Varken människa eller maskin kan dra tydliga slutsatser av vaga remissvar	41
	Referenser	43
	Bilaga 1 – Möjliggörande aktörer	45
	Bilaga 2 – Användarfeedbacken i sin helhet	48
	Bilaga 3 – Teknisk beskrivning av prototypen.....	65

Sammanfattning

Denna rapport beskriver Ekonomistyrningsverkets (ESV:s) och Tillväxtverkets arbete med att vidareutveckla prototypen Hitta remissvar, som lanserades i början av 2023. Hitta remissvar är en söktjänst som bygger på AI-baserade språkmodeller. Den ursprungliga prototypen hjälpte användaren att snabbt identifiera förslag som finns i statens offentliga utredningar (SOU) och departementsskrivelser (Ds), hitta liknande förslag i andra SOU:er och Ds:ar och se vilka remissinstanser som lämnat synpunkter.

Specifikt mål men omfattande syfte

Målet med projektet har varit att förbättra funktionaliteten i Hitta remissvar, sprida kunskap om tjänsten och utvärdera den. Ett större, mer indirekt, syfte har varit att visa på potentialen i vad som kan uppnås med språkbaseade AI-modeller tillämpade på offentlig text. Det långsiktiga målet är att lägga grunden för att utveckla AI-baserade analys- och beslutsmetoder som bidrar till bättre beslutsunderlag och en effektivisering av remisshanteringen i staten. Förhoppningen är också att exemplet Hitta remissvar ska bidra till en ökad medvetenhet om behovet av att förbättra databeredskapen för att kunna ta tillvara de effektivitets- och kvalitetsvinster som den nya tekniken möjliggör.

Denna rapport beskriver arbetet i projektet, sätter det i en kontext och tillgängliggör resultaten för en bredare målgrupp.

Många användarkontakter och ny funktionalitet

Under projektets gång har arbetsgruppen hämtat in synpunkter från användare genom en enkät, workshoppar, demonstrationer och möten.

Som ett resultat av projektet har prototypens funktionalitet uppdaterats och fler datakällor gjorts tillgängliga, så att det nu

- är möjligt att söka i utredningarnas konsekvensbeskrivningar
- sker en grafisk (och därmed mer exakt) identifiering av förslagsrutorna i utredningarna
- är möjligt att söka i remissvaren
- går att skapa matchningar mellan utredningsförslag och relevanta passager ur remissvar och konsekvensbeskrivningar
- sker en automatisk uppdatering av datakällorna
- genereras en unik URL för varje sökning, vilket gör det lätt att dela länkar till specifika sökningar.

Utöver detta har arbetsgruppen prövat olika typer av AI-stödda klassificeringar av remissvarens innehåll, uppdaterat gränssnittet och detaljerat beskrivit hur vi har hanterat de risker som uppstår när tidigare publicerade personuppgifter blir sökbara på nya sätt. Vi beskriver också några exempel på funktionalitet som skulle kunna förbättra Hitta remissvar ytterligare, men som vi inte har hunnit med att utveckla.

Möjligheter att permanenta och/eller bygga vidare på Hitta remissvar

För att ta Hitta remissvar från en prototyp till en permanent tjänst krävs ytterligare arbete, resurser och en organisation som har möjlighet att ta ett långsiktigt ägarskap. Det handlar om att säkerställa att tjänsten fungerar smidigt, uppfyller såväl tekniska som juridiska krav och kontinuerligt kan anpassas till förändringar i omvärlden.

Övergången från prototyp till en färdig tjänst innebär såväl initiala utvecklingskostnader för att färdigställa och anpassa tjänsten som löpande kostnader för att driva och förvalta tjänsten. Infrastrukturinköp kan också vara en engångskostnad som tillkommer. Även om prototypen skulle bli en permanent tjänst i någon myndighet så kvarstår arbete med att utveckla ytterligare funktioner som har efterfrågats av användare vi haft kontakt med. För att göra det enklare för andra att bygga vidare på projektets arbete vill vi att dokumenten ska finnas tillgängliga som öppna data och att vi ska kunna dela prototypens kod som öppen källkod. Förhoppningen är att kunna dela öppna data och källkod senare under 2024.

Lärdomar och rekommendationer för bättre nyttjande av textdata och AI-verktyg i offentlig sektor

Under arbetet med att ta fram och uppdatera Hitta remissvar har det blivit tydligt att det finns mycket övrigt att önska vad gäller databeredskapen i det offentliga Sverige. Här kan vi alla myndigheter bidra i än högre grad. Det behövs en ökad förståelse för att text är data, och att den behöver tillhandahållas i ett strukturerat och maskinläsbart format, som helst är möjlig att hämta via ett API. Först då går det att dra verklig nytta av den moderna AI-teknikens möjligheter att identifiera, analysera och generera text.

Andra lärdomar är att utvecklingen av AI är ett område där det sker snabba förändringar vad det gäller både tekniska och juridiska förutsättningar, och där vi på många sätt är beroende av vad andra aktörer gör, både inom och utom staten.

Vi ser att prototyper kan vara ett utmärkt komplement till utredningsverksamhet. Att ta fram något konkret kan, även om det inte är perfekt, skapa direkt nytta och göra att användarna kan skapa sig en bild av de möjligheter som användning av textdata och språkbaserade AI-modeller kan ge. En faktisk tjänst, om än i prototypform är ett illustrativt sätt att visa nyttor. Även små steg i innovationsprocesser tror vi leder till förbättringar som ytterst gagnar medborgare och samhället i stort.

1 Inledning

I detta kapitel redogör vi bl.a. för bakgrunden och projektets mål, syfte och genomförande.

1.1 Bakgrund

Varje år produceras hundratusentals sidor statlig utredningstext, och sammantaget lägger myndigheter och andra remissinstanser stora resurser på att lämna synpunkter på de remitterade utredningarna. Genom att myndigheter, organisationer och andra får ta del av förslagen kan de bli allsidigt belysta och konsekvenserna av förslagen bli klarlagda. Därmed ökar kvaliteten på regeringens och riksdagens arbete, till gagn för demokrati, rättssäkerhet och effektivitet.¹

Remissvar publiceras som pdf-filer med varierande struktur och datakvalitet på Regeringskansliets webbplats, [regeringen.se](https://www.regeringen.se), och ibland även på remissinstansernas webbplatser. Detta gör det tidskrävande att söka fram vad en viss remissinstans har tyckt tidigare i liknande frågor. Det har heller inte varit möjligt att söka igenom samtliga utredningar och remissvar efter vissa nyckelord eller samtliga svar från en särskild remissinstans.

1.1.1 Tidigare arbete och ursprunglig prototyp

Under slutet av 2022 början av 2023 utvecklade Ekonomistyrningsverket (ESV) prototypen Hitta remissvar i samarbete med Tillväxtverket, Statens servicecenter och Statskontoret. Utvecklingen bedrevs inom ramen för ESV:s datalabb. Prototypen är en söktjänst som bygger på AI-baserade språkmodeller och hjälper användaren att snabbt identifiera förslagen i statliga utredningar, hitta liknande förslag i andra utredningar och se vilka remissinstanser som har lämnat synpunkter. Hitta remissvar har funnits tillgänglig på datalabb.esv.se/remissai/ sedan mars 2023.

Data har hämtats från två separata källor. Utredningar (SOU och Ds) har laddats hem från riksdagens öppna data. Inlämnade remissvar har webbskrapats² från regeringens webbplats för remisser. Matchningarna mellan liknande förslag görs av Kungliga bibliotekets (KB:s) språkmodell KB-BERT som har tränats på stora mängder svensk text. En mer detaljerad beskrivning av prototypen finns i bilaga 3.³

1 SB PM 2021:1, Statsrådsberedningen, [Svara på remiss – om remisser av betänkanden och andra förslag från Regeringskansliet](#), s 4. Regeringskansliet.

2 Insamling av information direkt från webbsidor med hjälp av speciella program. Det kallas för skrapning eftersom programmen hämtar informationen från webbsidor och inte de bakomliggande databaserna. Definition hämtad från IDG, <https://it-ord.idg.se/ord/webbskrapning/>. Ändrad 15 november 2018. Hämtad 7 februari 2024.

3 För en beskrivning av ursprungsprototypen utifrån ett databeredskapsperspektiv, se rapporten Databeredskapsverkstaden (2023), [Databeredskap för språkteknologiska tillämpningar - Fallstudier från offentlig sektor](#), från s. 50.

Efter lanseringen i mars 2023 gick arbetet in i en ny fas, där ESV:s datalabb och Tillväxtverket inledde ett fortsatt samarbete för att sprida, utvärdera och förbättra funktionaliteten i den befintliga prototypen.

1.1.2 Snabb AI-utveckling i omvärlden

Merparten av arbetet med den första versionen av prototypen genomfördes före det stora genombrottet för generativ AI i form av lanseringen av ChatGPT. I och med lanseringen blev vissa tekniker som analys och generering av text och bild bättre och mer tillgängliga för applicerade tillämpningar. Sedan dess har utvecklingen inom AI-området gått fort, och verktygen har blivit mer lättillgängliga för fler.

Även om olika AI-verktyg har blivit mer tillgängliga för fler ändamål och användare, återstår viktiga skäl för myndigheter att skapa sina egna, skraddarsydda tjänster. Genom att anpassa AI-verktygen skapar myndigheterna möjligheten att mer exakt uppfylla sina specifika behov och samtidigt effektivisera hanteringen av stora mängder text från olika dokument och källor. En tjänst designad för ett särskilt ändamål, byggd på förprocesserad data från trygga källor, kan ge mer pålitliga svar och vara enklare för användaren att utnyttja än mer allmänna AI-verktyg. Det finns även flera fall där man vill hantera textdata på egna servrar snarare än att dela data i molntjänster och med externa parter.

Dessutom finns det en poäng för staten att lära sig om potentialen i AI och stora textdatamängder genom att experimentera med hur den nya tekniken kan användas och anpassas för att förenkla och effektivisera olika uppgifter. En del av de lärdomar som kan komma av ett sådant arbete är att myndigheternas sätt att tillhandahålla offentliga texter även påverkar andra aktörers möjligheter att utveckla och distribuera tjänster med AI-stöd, vilket kan öka transparensen och förenkla för användare av texterna. Läs mer om detta i avsnitt 6.1 om databeredskap.

1.1.3 Kopplingen till Tillväxtverket och ESV:s uppdrag

ESV ska enligt sin instruktion bl.a. utveckla och förvalta principer, regelverk och metoder för en effektiv ekonomisk styrning i staten och ge stöd till andra statliga myndigheter genom metoder, riktlinjer, utredningar, utbildning och rådgivning. ESV har också fått ett nytt uppdrag att stötta andra myndigheters och kommittéers arbete med konsekvensutredningar och konsekvensanalyser.⁴

Inom ramen för ESV:s datalabb, som är en avdelnings- och kompetensöverskridande funktion, utforskar ESV hur olika tekniska lösningar, inklusive AI-verktyg, kan

⁴ Förordningen (2016:1023) med instruktion för Ekonomistyrningsverket.

förbättra och effektivisera myndigheternas analyser och förståelse för data och i förlängningen bidra till bättre hushållning av statens resurser.

Tillväxtverket arbetar för hållbar tillväxt och konkurrenskraftiga företag i alla delar av Sverige. Effektiva och ändamålsenliga regler främjar företagens konkurrenskraft samt innovations- och omställningsförmåga. Tillväxtverket ska enligt sin instruktion bl.a. svara för

- metodutveckling, rådgivning och utbildning med anledning av förordningen (2007:1244) om konsekvensutredning vid regelgivning
- utveckla och föreslå nya insatser inom förenklingsarbetet
- främja en god kunskap hos myndigheter om hur företag påverkas av tillämpning av regelverk och förenklingsarbetet.⁵

ESV och Tillväxtverket ser gemensamt möjligheter med verktyg för att såväl effektivisera statens eget arbete med remissvar som det vidare behovet av att fler intressenter ska kunna bidra till bättre beslutsunderlag vid regelutvecklingsarbete.

1.1.4 Möjliggörande aktörer

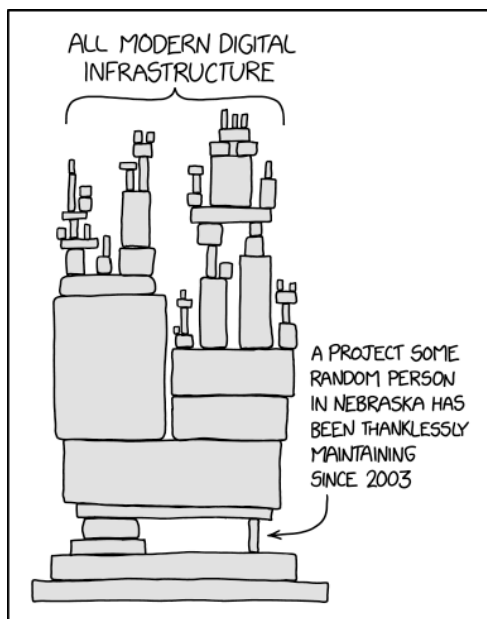
En utgångspunkt i vårt arbete är att remissvaren finns tillgängliggjorda på regeringen.se.

Hitta remissvar hade dock inte kunnat utvecklas utan de komponenter som tillhandahålls öppet av Kungliga Bibliotekets datalabb KB-labb (språkmodellen KB-BERT)⁶ och riksdagens öppna data (textdata från utredningarna). Inom ramen för projektet har arbetsgruppen träffat dessa möjliggörande aktörer för att uttrycka vår uppskattning för deras arbete, visa upp vad som kunnat göras tack vare dem samt inhämta synpunkter och utbyta erfarenheter.

I bilaga 1 redogör vi mer i detalj för dessa aktörers bidrag och arbete. En kort sammanfattning illustreras av figur 1, där blocket som den nedre pilen pekar på i denna kontext skulle kunna sägas motsvaras av riksdagens öppna data, och Hitta remissvar är någon av pytte-blocken överst. Både KB-labb, riksdagens öppna data och ESV:s datalabb har startats på myndigheternas egna initiativ.

⁵ Förordningen (2009:145) med instruktion för Tillväxtverket.

⁶ I vår applikation använder vi språkmodellen KB-BERT för att kunna matcha förslag mot varandra med hjälp av semantik. Språkmodellen är tränad på stora mängder svensk text. Se detaljerad beskrivning i Malmsten, M., Börjeson, L., och Haffenden, C. (2020), [Playing with Words at the National Library of Sweden--Making a Swedish BERT](#).

Figur 1: xkcd Dependency⁷

1.2 Projektets mål och syfte

Målet med projektet har varit att förbättra funktionaliteten i Hitta remissvar, sprida kunskap om tjänsten och utvärdera den. Syftet var att visa på potentialen i att använda AI-baserade språkmodeller i statsförvaltningen. Effektmålet på längre sikt var att utveckla AI-baserade analys- och beslutsmetoder som bidrar till bättre beslutsunderlag och en effektivare remisshantering i staten.

Utöver själva förbättringarna i prototypen skulle projektet utreda vilka steg som behöver tas för att åstadkomma en permanent tjänst samt använda och utvärdera flera språkmodeller.

Projektet syftar alltså till att åstadkomma förändring både direkt och handfast i det lilla (utvecklingen av prototypen) och mer indirekt genom påverkan på större system (det lärande och exempel vi bidrar med).

Både den nya versionen av Hitta remissvar och denna rapport är att betrakta som resultatet av arbetet och medel för att nå målen och syftet.

1.2.1 Genomförande och avgränsningar

Projektets arbetsgrupp bestod inledningsvis av två personer från ESV:s datalabb (en projektledare och en data scientist) och två personer från Tillväxtverket (som varit både behovsägare och resurser i projektet). Under hösten 2023 tillkom ytterligare en

⁷ Explain xkcd, 2347: *Dependency*, https://www.explainxkcd.com/wiki/index.php/2347:_Dependency. Uppdaterad 12 oktober 2023. Hämtad 7 februari 2024. CC-BY-NC

utredare från ESV och en datalingsvist från Tillväxtverket i projektet. Personerna i arbetsgruppen har deltagit i arbetet inom respektive organisations ordinarie uppdrag och budget. Tillväxtverket har bidragit med 300 000 kronor som möjliggjort timanställning av en student som hjälpt till med programmering och databearbetning, samt avrop av en konsult som bidragit i utvecklingen av användargränssnittet.

Projektets skulle huvudsakligen bedrivas under andra halvan av 2023. Insatserna i projektet behövde därför prioriteras utifrån vad som bedömdes vara möjligt att hinna med inom den tidsramen. Projektet inriktade sig på vidareutveckling direkt kopplad till den befintliga prototypen Hitta remissvar, och inkluderade därmed inte andra användarbehov som hade identifierats i anslutning till remissprocessen – t.ex. med fokus på remissammanställning, remittering eller författandet av betänkandet.

1.2.2 Kommunikation av resultatet

Projektet har bedrivits med utgångspunkten att resultaten ska delas öppet till externa intressenter. Det ligger i projektets syfte att den nytta och kunskap som skapats under projektets ska kunna användas för framtida innovation och utveckling efter att projektet har avslutats.

Denna rapport är basen för kommunikationen av projektets resultat och genomförande. Arbetsgruppen har använt generativ AI som stöd under delar av skrivandet. Följande aktiviteter ingår också i projektets kommunikationsplan:

- lanseringsseminarium av den uppdaterade prototypen
- utskick till alla som varit involverade i utvecklingen
- kontakter med Regeringskansliet
- försök att få ytterligare uppmärksamhet via pressmeddelande och en kortare artikelversion av slutrapporten som kommer att publiceras på ESV-forum.

1.3 Disposition och läsanvisning

I detta kapitel har vi redogjort för bakgrund och syfte med projektet. En fördjupad bakgrund som beskriver mer om de möjliggörande aktörerna KB-labb och riksdagens öppna data finns i bilaga 1.

I kapitel 2 sammanfattar vi den input arbetsgruppen fått via workshoppar, demonstrationer och enkät, samt vad vi vet om den faktiska användningen av prototypen. En mer utförlig redogörelse finns i bilaga 2.

I kapitel 3 beskriver vi de insatser som har gjorts för att utveckla prototypens funktionalitet. Där finns även en ingående beskrivning av hur vi har hanterat de nya risker som uppstår för hanteringen av personuppgifter när befintlig information blir

sökbar på nya sätt. Kapitlet avslutas med en lista över funktioner som vi hade velat utveckla, men som inte rymdes inom projektets tid, inriktning, mandat eller resurser.

Kapitel 4 redogör för vilka steg och kostnader som ytterligare krävs för att kunna ta prototypen till en färdig tjänst, både för den initiala utvecklingen och den långsiktiga driften och det långsiktiga underhållet.

I kapitel 5 beskriver vi dels värdet av öppna data och öppen källkod, dels varför arbetsgruppen trots våra ambitioner inte har lyckats dela detta från vårt projekt (ännu). I väntan på att komma i mål i den delen finns det i bilaga 3 en mer teknisk beskrivning av prototypen som kan fungera som inspiration/instruktion för utvecklare som vill bygga vidare på det vi har gjort. Det finns även en fördjupad beskrivning av den ursprungliga prototypens funktion ur ett databeredskapsperspektiv i Databeredskapsverkstadens rapport *Databeredskap för språkteknologiska tillämpningar – Fallstudier från offentlig sektor*⁸.

I rapportens sista kapitel, kapitel 6, redogör vi för arbetsgruppens lärdomar och rekommendationer för att ta tillvara på den potential som textdata och AI-verktyg skulle kunna leda till. Det huvudsakliga budskapet är att det offentliga måste öka förståelsen för att text är data och att det behövs åtgärder både för att tillgängliggöra befintliga texter i maskinläsbara format och att framtidssäkra kommande textdata så att det blir rätt från början i framtiden.

För dig som vill bygga något liknande rekommenderar vi främst kapitel 2 och 3, användarfeedbacken i bilaga 2 och tekniska beskrivningen i bilaga 3. För dig som är intresserad av vad som krävs för utveckling av databeredskap och AI i offentlig förvaltning rekommenderar vi framförallt kapitel 4,5 och 6 samt bilaga 1.

⁸ Databeredskapsverkstaden (2023), [Databeredskap för språkteknologiska tillämpningar - Fallstudier från offentlig sektor](#), från s. 50.

2 Utvärdering, testning och användarfeedback

I detta kapitel beskriver vi kortfattat den återkoppling arbetsgruppen fått genom enkät, demonstrationer och workshoppar, samt redogör för vad vi vet om den faktiska användningen av Hitta remissvar. Utifrån detta kan vi också uttala oss om den potentiella nyttan av verktyget, samt möjliga förbättringsåtgärder. En mer detaljerad beskrivning av användarfeedback finns i bilaga 2.

2.1 Demonstrationer

En mängd olika demonstrationer och presentationer av Hitta remissvar genomfördes under projektets gång. Deltagarna var ofta positiva till initiativet och applikationen, som visar på möjligheterna med AI och språkteknologi inom staten. Men det framfördes också flera synpunkter och önskemål på särskilda typer av förändringar eller funktioner.

Bland annat framkom önskemål om att kunna sortera sökresultat på ett annorlunda sätt. Flera deltagare ville också få möjligheten att gruppera resultaten efter den utredning det gällde i stället för att visa varje enskilt förslag för sig. Något annat som ofta togs upp var att man gärna ville kunna söka bland äldre remissvar än från 2016.⁹ Ytterligare ett förbättringsområde som pekades ut var att det saknas hantering av myndigheter vars namn har förändrats genom åren, eftersom myndighetsnamnen extraheras från filnamnen och inte hanteras som objekt.

2.2 Enkät

I anslutning till applikationen på datalabb.esv.se kan användarna svara på en enkät om sin användarupplevelse. Enkäten består av tio frågor, varav fem tillåter fritextsvar. Arbetsgruppen har sammanställt resultatet av enkäten utifrån de 39 tillgängliga svar som hade lämnats vid tidpunkten för skrivandet av den här rapporten. Genom enkäten har arbetsgruppen fått många förslag på möjliga förbättringar för Hitta remissvar, samt fått veta mer om applikationens användare.

De huvudsakliga resultaten från enkäten om hur Hitta remissvar används:

- Besvara remisser är den vanligaste arbetsuppgiften som tjänsten används för.
- Medelvärdet för hur användbar tjänsten är för några exempel på arbetsuppgifter (bl.a. ”besvara remisser” och ”bereda beslutsunderlag”) blev 6,3 på skalan 0–10.
- Att hitta underlag som annars inte hade identifierats angavs som den största nyttan med tjänsten.

⁹ På regeringen.se finns remissvar från 2016 publicerade. Vi har inte haft möjlighet att samla in remissvar på annat sätt än skrapning av den webbplatsen.

- De flesta användare har svårt att uppskatta hur mycket tid de sparar på att använda tjänsten. Förutom ”vet ej” svarar de flesta att de kan spara 1–2 timmar i månaden.
- Användare ser flera risker med att använda tjänsten. Två exempel på risker:
 - Begränsningen för vilka remissvar som finns att hitta i tjänsten. En person nämner problematiken med att förlita sig på information man kanske bara läst ett utdrag ifrån (det utdrag som presenteras i applikationen).
 - Man kan luras att tro att det inte finns några relevanta resultat på ens sökning på grund av att vissa dokument inte finns med i applikationen, eller på grund av att sökningen hade kunnat göras på ett annorlunda sätt.
- Att kunna söka i och sortera ut delar i remissvar som handlar om ett specifikt förslag är den ytterligare funktion som anses mest användbar. Denna funktion fanns inte i versionen av applikationen som var aktuell när enkätsvaren sammanställdes för denna rapport.
- Övriga kommentarer innefattade bl.a. önskan om ett API, att förslag från samma utredningar sorteras eller visas ihop, och att kunna se hur stor procent av remissvaren från en särskild utredning som inte innehåller några synpunkter alls.

De huvudsakliga resultaten från enkäten om användarna av Hitta remissvar:

- Den typiska användaren svarar på fem eller fler remisser om året.
- Den största delen av användarna har jobbat med utredningar och remisser i mer än 5 år.
- Nästan dubbelt så många kvinnor som män svarade på enkäten.

Materialet är för litet för att kunna dra statistiskt säkerställda slutsatser och generalisera om skillnader utifrån kön. Men bland de som besvarat enkäten kan vi se att det är ungefär dubbelt så stor andel kvinnor än män som anger att prototypen kan ge dem perspektiv och infallsvinklar de inte tänkt på innan, och att prototypen kan bidra till att resultatet av deras arbete blir bättre. Kvinnorna som svarat på enkäten är också generellt mer positivt inställda än männen till ytterligare funktioner.

2.3 Testning av klassificering och matchning med hjälp av AI

Arbetsgruppen har genomfört tre workshoppar för att få återkoppling från användare och viss kvalitativ utvärdering av möjliga funktioner i prototypen. Sådana möjliga funktioner är t.ex. automatisk bortsortering av remissvar helt utan synpunkter och att kunna sortera svaren utifrån om de är positiva, negativa eller neutrala till utredningens förslag. Att klassificera text baserat på sentiment (känsla) kallas för göra en sentimentanalys.

2.3.1 Klassificering, sentimentanalys och sammanfattningar med hjälp av generativ AI

Under workshop 1 och 2 testade arbetsgruppen att med generativa språkmodeller utföra flera olika uppgifter. Dessa var klassificering av om remissvar innehåller synpunkter eller ej samt sammanfattning av remissvar. Under workshop 1 testade vi även att använda språkmodellerna för sentimentanalys.

Dessa uppgifter utfördes med hjälp av två generativa språkmodeller: GPT-4 och GPT-SW3. Arbetsgruppen valde att testa att använda generativa språkmodeller eftersom det är en teknik som har utvecklats snabbt på senare tid och fått stort genomslag i samhället. En s.k. storskalig språkmodell, som GPT-4 eller GPT-SW3, har tränats på mycket stora textmängder och är tänkta att kunna fungera som en generell lösning för olika typer av texthanteringsuppgifter inom vitt skilda områden. Därför ville vi utforska generativa språkmodellens potential för att hantera remissvar.

Klassificeringstyp 1: Sentimentanalys

Sentimentanalys innebär att klassificera ett dokument baserat på dess sentiment (känsla). De vanligaste klasserna att använda sig av är positivt, negativt och neutralt. Det finns flera möjliga metoder för denna typ av klassificering såsom lexikal sentimentanalys¹⁰ eller neurala nätverk¹¹.

Klassificeringen gick till på så vis att vi gav GPT-modellerna enkla instruktioner om att bedöma ett remissvar som positivt, negativt eller neutralt. Språkmodellerna fick följande instruktion (prompt):

Du kommer att få text från ett remissvar och sen ett förslag från en statlig utredning. Ditt jobb är att avgöra om remissvarstexten är positivt, negativt eller neutralt inställd till förslaget.

Det här är ett utdrag ur remissvaret från {*remissinstans*}:
{*synpunkt*}

Vilken åsikt har {*remissinstans*} i texten ovan på innehållet i följande förslag?

{*förslag*}

Svara positiv, negativ eller neutral.

¹⁰ Lexikal sentimentanalys är en sentimentanalys som baseras på vilka ord som finns med i en text, t.ex. "Jag är glad" innehåller 1 positivt ord och inget negativt ord, "Jag är glad och ledsen" innehåller 1 positivt och 1 negativt ord.

¹¹ Så kallade artificiella neurala nätverk är en teknik som används inom maskininläring/AI. Med hjälp av dessa kan man "träna" modell på data med positiva, negativa och neutrala meningar och utvärdera resultatet. Om modellen fungerar bra kan den användas för att sentimentklassificera ny text.

2.3.1.1 Klassificeringstyp 2: Synpunkt eller inga synpunkter

För varje remissvar ville arbetsgruppen kunna göra en automatisk bedömning av ifall svaret innehåller synpunkter eller inga synpunkter om ett särskilt förslag. Vi använde oss av utdrag ur remissvaren och bad modellerna att svara på en enkel och tydlig fråga: handlar detta stycke om detta remissvar?

Flera olika instruktioner (*prompts*) testades för att få ett tillfredställande resultat. Båda språkmodellerna fick varianter av följande instruktion:

Du kommer att få text från ett remissvar och sen ett förslag från en statlig utredning. Ditt jobb är att avgöra om remissvarstexten innehåller eller inte innehåller åsikter om utredningsförslaget.

Det här är ett utdrag ur remissvaret från {remissinstans}:
{utdrag}

Uttrycker {remissinstans} i texten ovan några synpunkter på innehållet i följande förslag?
{förslag}

Svara ja eller nej. Tänk på att det mest sannolika är att texten innehåller synpunkter om utredningen som inte är relaterade till det specifika förslaget. I de fallen ska du svara nej.

2.3.1.2 Sammanfattning av remissvar

Arbetsgruppen bad GPT-4 och GPT-SW3 att sammanfatta remissvaren. Prompten såg ut så här:

Du kommer att få text från ett remissvar och sen ett förslag från en statlig utredning. Ditt jobb är att sammanfatta vilka synpunkter remissvaret har på förslaget.

Det här är ett utdrag ur remissvaret från {remissinstans}:
{synpunkt}

Uttrycker {remissinstans} i texten ovan några synpunkter på följande förslag?
{förslag}

Tänk på att texten också kan innehålla andra synpunkter som inte hör till detta förslag. Sådana synpunkter ska du inte ta med.

2.3.1.3 Resultat för workshop 1 och 2

Att använda generativa språkmodeller för sentimentanalys visade sig vara en utmaning. GPT-SW3 gav nästintill alltid en positiv sentimentklassificering, trots att flera av de bedömda textstyckena bedömdes som negativa när de granskades av en person. GPT-4 presterade något bättre.

Efter den första workshopen gjorde arbetsgruppen bedömningen att sentimentklassificering med hjälp av generativa modeller inte skulle undersökas vidare inom projektet, eftersom försöken inte gav särskilt goda resultat. Dessutom är uppgiften även svår att utföra manuellt vilket gör att det blir svårt att utvärdera en eventuell automatisk sentimentanalys.

Då uppgiften som sådan därmed bedömdes som för komplex i relation till projektets tid och resurser har arbetsgruppen inte heller närmare undersökt andra metoder för uppgiften, såsom lexikal sentimentanalys. I och med detta resultat och denna bedömning valde arbetsgruppen att inte prioritera vidare undersökning och utveckling av sentimentanalys som funktion.

Båda språkmodellerna fungerade däremot generellt bra för att bedöma om ett remissvar innehöll synpunkter om utredningen eller ej. GPT-SW3 hade dock svårt att följa instruktionen, men visade bättre resultat när man manuellt angav synpunkten före förslaget. Denna metod kan dock bli kostsam och klassificering är också en uppgift där andra typer av språkmodeller än generativa generellt kan prestera väldigt bra. Därför bedömde arbetsgruppen att alternativa metoder borde undersökas om en liknande funktion ska utvecklas.

Ett möjligt alternativ är en regelstyrd metod som bedömer ett remissvar som med eller utan synpunkt utifrån en eller flera regler. Reglerna kan exempelvis vara baserade på innehåll (t.ex. finns frasen "inga synpunkter" eller dylikt i remissvaret) eller/i kombination med krav på längd på remissvaret (t.ex. om dokumentet innehåller färre än X antal ord). En annan metod är att träna en egen klassificerare med hjälp av träningsdata i form av redan klassificerade remissvar vars klassificering har kontrollerats manuellt.

GPT-4 fungerade mycket väl för sammanfattningar, medan GPT-SW3 generellt gav sämre sammanfattningar. GPT-SW3 gav dessutom väldigt olika resultat för varje försök trots identiska instruktioner, vilket i sig inte är bra då konsekventa resultat är att föredra.

Möjligheten att utveckla funktioner i Hitta remissvar som baseras på generativa språkmodeller beror även på ett antal andra faktorer utöver vilken användarnytta arbetsgruppen bedömer att de kan medföra. Utvecklingen av storskaliga

språkmodeller och liknande språkteknik går mycket snabbt. Språkmodellerna tillhandahålls också både som kommersiella produkter, öppna lösningar och offentligt finansierad forskningsinfrastruktur samt olika kombinationer av dessa. Under projektiden bedömde vi att rättsläget, leverantörsförhållanden och kostnader var så osäkra att det inte var möjligt att bygga in ett beroende till någon av dessa modeller/produkter i applikationen.

2.3.2 Matchningar mellan förslag och remissvar

Under den tredje workshopen undersöktes en annan typ av funktion kopplad till remissvarens innehåll, nämligen att kunna matcha ett förslag i en remiss mot utdrag från remissvar. Användaren kan antingen välja att söka bland alla remissens remissvar, eller söka i utvalda remissvar och hitta det mest relevanta stycket ur dessa. För att matcha förslag och remissvar har vi använt oss av en semantisk sökfunktion (se bilaga 3 för detaljer).

2.3.2.1 Resultat workshop 3

Den undersökta funktionen gav generellt goda resultat, men med vissa undantag. Det är vanligt att remissinstanserna citerar förslagen. Utdrag med dessa citeringar hamnar ofta högt upp i rankningen av matchande utdrag (till ett förslag). Men i de fall där remissinstanserna har återgett förslaget utan att nämna sin egen inställning till det blir matchningen inte relevant för användaren.

Ibland är det också svårt för användarna att bedöma hur väl rankningen fungerar. Men under workshopen bedömdes rankningen i flera fall som inte fullt ut korrekt. I ett fall hade själva utredningen i fråga publicerats som ett remissvar på regeringen.se. Detta medför att utdrag ur utredningen matchas mot identiska utdrag från det felaktiga "remissvaret".

2.4 Förenkling av användargränssnitt

En UX-designer fick i uppgift att ge förslag på hur gränssnittet skulle kunna göras mer användarvänligt. Nedan följer ett urval av förslagen:

- Använd en sökruta i stället för tre.
- Visa instruktioner för hur sökningen går till på applikationens startsida.
- Texten som beskriver applikationen bör vara kortare och mer koncis.
- Använd en rullgardinsmeny för remissinstanser i stället för en sökruta.

Fler förslag från UX-designern samt mer detaljer om förslagen finns i bilaga 2.

2.5 Summering av återkommande synpunkter

Här följer en summering av de vanligaste synpunkterna arbetsgruppen fick från enkät, demonstrationer och workshoppar:

- Användarna vill gärna se utökade data för remissvar (just nu finns inga remissvar från tidigare år än 2016).
- Många användare vill kunna gruppera sökresultaten efter utredning i stället för efter förslag.
- Remissinstansernas namn kan variera trots att namnen syftar på en och samma instans. Användarna önskar att ett och samma namn för instanserna används för hela Hitta remissvar.
- Att kunna söka i och sortera ut delar av remissvar är den ytterligare funktion som anses mest användbar av användarna. Denna funktion har prioriterats och finns i den nya versionen av Hitta remissvar.

2.6 Faktisk användning av Hitta remissvar

I detta avsnitt beskriver vi vad arbetsgruppen känner till om i vilken utsträckning som Hitta remissvar används.

2.6.1 Besöksstatistik

Sedan hösten 2023 har vi samlat in besöksstatistik för Hitta Remissvar. Bara besökare som godkänner kakor räknas i statistiken. Därför är det svårt att veta hur stor andel av besöken som statistiken har lyckats fånga¹², men om man antar att det under hösten 2023 är en ungefär lika hög andel som godkänner kakor för Hitta remissvar så kan siffrorna i alla fall säga något om trenden. Tabell 1 visar hur det ser ut på månadsbasis.

Tabell 1. Antal besök per månad (av personer som accepterat kakor)

Månad	Antal besök
September 2023	163
Oktober 2023	150
November 2023	182
December 2023	108

På veckobasis har Hitta remissvar oftast 40–50 besök, men med vissa dippar neråt 20 besök var tredje till fjärde vecka. En tydlig nedgång ses också vid jul och nyår, vilket åtminstone delvis förklarar de låga siffrorna för december i tabell 1.

¹² Enligt Internetstiftelsens undersökning från 2022 är det ungefär en fjärdedel av internetanvändarna som väljer bort cookies. Internetstiftelsen, *Var fjärde internetanvändare väljer bort cookies*, <https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2022/vad-begransar-var-internetanvandning/var-fjarde-internetanvandare-valjer-bort-cookies/#vilka-klickar-ur-cookies>. Hämtad 8 februari 2024.

Figur 2. Graf över antal besök på Hitta remissvar på veckobasis

2.6.2 Hitta remissvar som en del av remissvarsrutinen

Pensionsmyndigheten och Tillväxtverket har uppdaterat sina interna riktlinjer för att svara på remiss med en uppmaning att använda Hitta remissvar som en del i utredningsarbetet. Tillsammans med besöksstatistiken är det en indikation på att applikationen används.

3 Prototypens nya funktionalitet

Utifrån arbetsgruppens egna idéer, tester och den feedback användare lämnat bestämde vi oss för att göra ett antal åtgärder som skulle kunna förbättra och utvidga funktionaliteten för Hitta remissvar. I det här kapitlet beskriver vi de nya funktionerna i den uppdaterade prototypen och hur utvecklingsarbetet har gått till. Vi beskriver också sådan utveckling som vi hade velat göra men inte haft möjlighet att genomföra inom projektets ram. En mer detaljerad beskrivning av tekniska detaljer finns i bilaga 3. Den nya prototypen lanseras 21 mars 2024, d.v.s. något efter publiceringen av denna rapport.

3.1 Grafisk identifiering av förslagsrutor för ökad exakthet

I den ursprungliga prototypen Hitta remissvar identifierades förslagen med hjälp av en enkel sökning på ”förslag:”, då det visat sig fånga de flesta förslagen i utredningarna. Vi sorterade bort några vanliga överfångster och exkluderade innehållsförteckningar. Sedan använde vi den text som finns ca 300 ord efter texten ”förslag:”, men med vissa brytregler (t.ex. om det kommer ett nytt ”förslag:” eller en ny rubrik). Med denna metod är det dock svårt att identifiera när förslagen slutar och vissa förslag saknar ordet “förslag:”, vilket gör att metoden missar dessa helt.

I den nya prototypen ville arbetsgruppen få en mer exakt extraktion av förslagen genom att med hjälp av grafiska metoder identifiera, och sedan presentera och söka på, enbart det som står inom den markerade förslagsrutan. Sådana rutor omringar förslagen i de allra flesta av utredningarna. Med hjälp av bl.a. mjukvaran OpenCV har vi identifierat alla tecknade former, urskilt vilka som kan vara förslagsrutor och extraherat texten som finns i varje ruta.

Då det förekommer rutor i utredningarna där texten inte är förslag har vi i den uppdaterade prototypen kompletterat denna grafiska identifikation med vissa enkla regler av det slag som användes för förslagsextraktion i första versionen. Till exempel inkluderar vi alla rutor som innehåller “förslag:”, men exkluderar rutor som innehåller “bedömning:”, “faktaruta:”, “slutsats:” eller “överväganden:” – om de inte samtidigt innehåller “förslag:”.

Den nya metoden resulterade i att tjänsten numera identifierar över 2 500 fler förslag och filtrerar bort ungefär 1 000 träffar som inte var förslag. Dessutom blir kvaliteten bättre eftersom tjänsten nu enbart återger förslagets exakta text och inte överfångsten efter att förslagsrutan slutar. Även om den nya metoden ger vissa falska positiva resultat (säger felaktigt att något är ett förslag) och falska negativa resultat (missar förslag) som den tidigare metoden fångade, är vår bedömning att den nya metoden är mer träffsäker.

Vi har också parsat¹³ strukturen i utredningarna på ett bättre sätt så att vi får ut underrubriksstrukturen, så att denna metadata finns för varje förslag.

Dessa förändringar ger också förutsättningar för mer relevanta sök- och matchningsfunktioner. Men arbetsgruppen har inte utvärderat om sökningarna blir bättre med eller utan kontexten som följer utanför rutan (post-text) eller med eller utan rubrikstrukturen. Det hade varit intressant att göra om vi haft mer tid. Den nya prototypen visar dock ett mer exakt utdrag ur förslagsrutan från utredningen, men post-texten används vid matchningarna.

3.2 Göra remissvaren sökbara

Den ursprungliga prototypen möjliggjorde bara sökningar i utredningens förslag. Arbetsgruppen ville utöka funktionaliteten genom att möjliggöra sökning också i remissvaren, vilket kräver nya kopplingar mellan datakällan och vårt gränssnitt, samt i vissa fall processering av remissvar med sämre maskinläsbarhet.

Till skillnad från statliga utredningar som följer ungefär samma mall kommer remissvaren från tusentals olika organisationer, som använder många olika format. Vissa svar är tiotals sidor långa och har en tydlig struktur medan andra består av några enstaka meningar eller flera stycken utan rubriker. Ytterligare några är inskannade fullmäktigeprotokoll eller andra bilddokument.

I den nya versionen är remissvaren uppdelade i stycken (segment). Dessa stycken indexeras (dvs. läggs in i sökmotorn) på samma sätt som utredningsförslagen. Det gör det möjligt att göra fritextsökningar i remissvarens text och få träff på segment ur remissvaren.

3.2.1 Hantera risker med personuppgifter

Alla verksamheter som hanterar personuppgifter måste följa dataskyddsförordningen (GDPR). Reglerna grundar sig i de mänskliga rättigheterna och handlar om att alla människor har rätt till skydd av sina personuppgifter. Några av de grundläggande principerna i dataskyddsförordningen är att bara samla in personuppgifter för specifika, särskilt angivna och berättigade ändamål och att inte behandla fler personuppgifter än vad som behövs för ändamålen.

¹³ Parsa, från engelskans parsing, innebär att dela upp och analysera i mindre delar. Exakt betydelse beror på sammanhang, men innebär i vårt fall att vi delat upp utredningstexten i kortare bitar och använt reguljära uttryck och metadata (t.ex. fontstorlek, fetstil och position) för att härleda vad som är rubriker.

Personuppgifter i Hitta remissvar

För att förstå hur tjänsten används samlar Hitta remissvar in ett fåtal uppgifter om de användare som godkänt det med hjälp av webbanalysverktyget Matomo. Det gäller bl.a. IP-nummer som betraktas som en personuppgift.

Men tjänsten behandlar också personuppgifter som förekommer i dokumenten som bearbetas. För statliga utredningar handlar det främst om personnamn för dem som har varit del av kommittéer, expertgrupper och/eller utredningssekretariat. För remissvar handlar det om personnamn och ibland kontaktuppgifter för dem som har handlagt eller beslutat ärendet. För kommuners remissvar är det inte ovanligt att det följer med hela protokoll från fullmäktige, inklusive namn på alla som deltagit.

Det förekommer även namn på forskare och experter vars produktion används som underlag till dokumenten. I vissa fall kan en bilaga till ett remissvar innehålla ytterligare personuppgifter som inte är relevanta för tjänstens funktion.

Alla dokument som används i Hitta remissvar hämtas från de offentliga källorna regeringen.se och riksdagens öppna data. Men i den mån det förekommer personuppgifter i dokumenten är hämtningen och samlandet av dessa i sig en hantering av personuppgifter.

När remissvar blir sökbara skapas ny sorts åtkomst till befintliga personuppgifter

I den tidigare prototypen kunde användarna bara söka i förslagsrutorna, där i princip inga personuppgifter förekommer. Men i den nya versionen går det att söka i remissvarens text, vilket skapar en ny sorts åtkomst till personuppgifter som finns i dem. Det skulle då kunna gå att söka på en persons namn och få träff på samtliga remissvar som personen omnämns i. På så sätt skulle det kunna vara möjligt att koppla en viss individ till en eller flera organisationer och att kartlägga en persons roller i dessa organisationer över tid. Som regel redogör dock remissvar för en organisations ståndpunkt och inte en individs. Men det kan vara svårt att skilja dessa åsikter åt, särskilt i mindre organisationer med få medlemmar eller anställda.

Arbetsgruppen bedömer att prototypen gör det lättare att koppla samman individer och åsikter och kartlägga dem. Vidare bedömer vi inte att personuppgifterna behövs för att uppnå ändamålet med prototypen. Mot den bakgrunden har vi valt att ta bort möjligheten att söka efter personnamn och andra identifierande personuppgifter såsom telefonnummer och e-postadresser i tjänsten. Det effektivaste sättet att göra detta på är att identifiera uppgifterna och ta bort dem från textmängderna som indexerats, dvs. textmängderna som läggs in i sökmotorn och blir sökbara.

Åtgärder för att minska sökbarheten för personuppgifter

Det är lätt att hitta telefonnummer och e-postadresser med hjälp av reguljära uttryck (s.k. *regex*). En analys som arbetsgruppen genomförde efter rensningen visar att denna metod är väldigt träffsäker och endast missar några enstaka fall där uppgiften blir felformaterad. Till exempel att “0” (noll) blir “O” under den optiska teckenigenkänningsprocessen (*OCR*) som görs i datahanteringen som vi använder för att få ut texter ur bildfiler med inskannade remissvar.

Tabell 2. Två reguljära uttryck som används för att fånga telefonnummer och e-postadresser i dokumentens text

Reguljärt uttryck	Exempel
<code>\+46\s?\d{1,2}[s-]?\d{3}s?\d{2}s?\d{2}[s-]?\d{3}s?\d{2}s?\d{2}</code>	0724 808 808 +46 724 808808
<code>\b[A-Za-z0-9._%+-]+@[A-Za-z0-9.-]+\.[A-Za-z]{2,}\b</code>	förnamn.efternamn@esv.se kontakt@efternamn.com

Att hitta personnamn är betydligt svårare. En lösning skulle kunna vara att använda sig av en lista och Statistiska centralbyrån (SCB) publicerade fram till förra året en sådan. Den innehåller dock endast de icke-unika namn som finns i folkbokföringsregistret i Sverige¹⁴ (ungefär 100 000 tilltalsnamn). Om man summerar alla namnbärare och jämför siffran med Sveriges befolkning kan man se att det sannolikt finns över 200 000 unika namn som inte finns i listan. Att leta efter namn i texten skulle dessutom leda till för många falska positiva resultat, eftersom många namn förekommer i texter som vanliga ord (t.ex. Bo, Per, Lena, Sten) eller som en del av fraser som inte är personuppgifter (t.ex. Olof Palmes gata, Raoul Wallenbergs stiftelse eller Anna Lindhs park).

För att kringgå denna svårighet använde arbetsgruppen tekniken **namnigenkänning**, dvs. en maskinlärningsmodell som har tränats specifikt för att annotera text och klassificera dessa ord. Vi använde modellen bert-base-swedish-cased-ner¹⁵, som har tagits fram av Kungliga Bibliotekets KB-labb med en svensk träningsdatamängd. För att säkerställa att modellen är träffsäker gjorde vi en särskild analys över alla remissvar från 2017.

Modellen användes för att ta bort 10 504 personnamn från remissvaren. Därefter gjorde vi ett automatiserat sök av de 1 000 populäraste namnen i SCB:s lista. Vi hittade 122 kvarvarande namn. Det betyder att drygt 1 procent av personnamnen inte identifierades av modellen.

¹⁴ Statistiska centralbyrån, Namnstatistik, <https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/amnesovergripande-statistik/namnstatistik/>. Hämtad 30 januari 2024.

¹⁵ KB-labb på HuggingFace, *Swedish BERT Models*, <https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/amnesovergripande-statistik/namnstatistik/>. Hämtad 30 januari 2024.

Efter en genomgång av 10 procent av de 10 504 identifierade personnamnen upptäckte vi även att 20 av dem inte tillhörde personer. Dessa var t.ex. organisationer (“Boko Haram”, “C More”) eller platser (“Giron Sámi Teáhter”). Om detta urval är representativt innebär det ungefär 2 procent av falska positiva.

En kvalitativ analys av namnen som blev kvar eller felaktigt togs bort visade att en del av dessa namn hanterades fel eftersom de var felformaterade. Bland resten hittades inget tydligt mönster som skulle kunna tyda på att modellen diskriminerar vissa kategorier av personer.

I det här avsnittet har vi beskrivit hur arbetsgruppen jobbat för att minimera riskerna i vår personuppgiftsbehandling. En annan lösning skulle kunna vara att kräva att användarna registrerar sig, formulera användarvillkor för hur data får användas, och därmed kunna spåra om någon har använt databasen på ett otillåtet sätt.

Riksdagens öppna data har användningsvillkor som specificerar hur data får användas, även om de inte kräver registrering. I användarvillkoren framgår bl.a. att “den som använder riksdagens öppna data måste försäkra sig om att behandling av personuppgifter sker i enlighet med EU:s allmänna dataskyddsförordning och lagen (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning”.¹⁶ Dessa användarvillkor är ytterligare ett av skälen till att vi behöver ha en medveten hantering av personuppgifter.

3.3 Möjliggöra sökning i utredningarnas konsekvensbeskrivningar

I 14–16 §§ kommittéförordningen (1998:1474), som statens offentliga utredningar förväntas följa, finns krav på att göra konsekvensbeskrivningar. I 15 § beskrivs ett antal perspektiv som ska beaktas i utredningen, däribland konsekvenser för kommunala självstyret, brottsligheten, företaget och jämställdheten. Ifall förslaget innebär nya eller ändrade regler hänvisar kommittéförordningen även till 6 och 7 §§ i förordningen (2007:1244) om konsekvensutredning vid regelgivning.

Eftersom både ESV och Tillväxtverket har särskilda roller i att stötta kommittéernas arbete med konsekvensutredningar var det intressant för oss att utveckla funktioner som stödjer detta arbete.

Genom att extrahera och möjliggöra sökning i utredningarnas konsekvensbeskrivningar ville arbetsgruppen ge användaren möjlighet att se om det finns förslag/utredningar som bedöms ha en viss typ av konsekvens. Ursprungligen var idén att både erbjuda fritextsökning och filtrering utifrån de konsekvenser som ska

¹⁶ Sveriges riksdag, [Användningsvillkor](https://www.riksdagen.se/sv/dokument-och-lagar/riksdagens-oppna-data/avandarstod/avandningsvillkor/), <https://www.riksdagen.se/sv/dokument-och-lagar/riksdagens-oppna-data/avandarstod/avandningsvillkor/>. Hämtad 30 januari 2024.

bedömas enligt nu gällande kommittéförordning och konsekvensutredningsförordning. Filtrerings-funktionen lade vi dock på is. Dels på grund av tids- och komplexitetsskäl, dels på grund av att det under projektetiden fortfarande var osäkert om några specifika konsekvenskategorier kommer att nämnas i den föreslagna nya konsekvensutredningsförordningen och uppdaterade kommittéförordningen.

För att extrahera konsekvensskrivningar ur utredningarna har vi sökt efter kapitelrubriker och underrubriker som innehåller "konsekvens" eller "följder". Vi har sedan tagit ut all text under sådana rubriker, delat upp den i kortare bitar och gjort dessa sökbara i tjänsten. Detta förenklade tillvägagångssätt gör att vi missar eventuella konsekvensbeskrivningar som står insprängda under andra kapitel och rubriker. Arbetsgruppen bedömer dock att merparten av utredningarna har konsekvensbeskrivningen samlad i ett utpekade kapitel.

3.4 Matchningar mellan förslag, remissvar och konsekvensutredningar

I den ursprungliga prototypen matchades förslag mot de semantiskt mest liknande förslagen. Nu när fler datatyper har tillgängliggjorts i prototypen skapas också möjligheten till fler sorters matchningar, både inom samma datatyp och mellan datatyper. Totalt finns nio olika sätt att söka i informationen.

3.4.1 Matchningar inom samma datatyp

Eftersom prototypen nu innehåller tre olika datatyper finns följande tre olika möjligheter att hitta liknande förslag av samma typ:

- Vilka förslag som är mest lika ett valt förslag (prototypens ursprungliga funktionalitet).
- Vilka utdrag ur ett remissvar som är mest lika ett valt utdrag ur ett remissvar.
- Vilka utdrag ur en konsekvensbeskrivning som är mest likt ett valt utdrag ur en viss konsekvensbeskrivning.

Inom samma datatyp har arbetsgruppen valt att kalla sökningen för "visa liknande förslag/remissvar/konsekvenser".

3.4.2 Matchningar mellan olika datatyper

Både remissvar och konsekvensbeskrivningar blir uppdelade i mindre stycken och indexeras. För förslagen indexeras allt som står i förslagsrutan. De genererade inbäddningarna kan sedan jämföras med varandra. För varje datatyp (förslag eller utdrag ur remissvar eller konsekvensutredningar) genereras en lista av träffar ur respektive annan datatyp som är semantiskt närmast. Det är dessa listor som kommer fram när användaren trycker på "matchande förslag", "matchande remissvar" eller "matchande konsekvenser".

Vilka möjligheter detta ger beskriver vi kortfattat i de tre följande avsnitten.

Matchningar mellan förslag och relevanta passager ur remissvar

Många remissvar kommenterar bara ett urval av utredningens förslag, utifrån vad som är remissinstansens expertis- eller intresseområde. Därför är det många remissvar som inte behandlar just det förslag användaren är intresserad av att läsa synpunkter på.

För att användarna ska kunna få se vilka remissvar (och vilken del av remissvaret) som gäller ett specifikt förslag utvecklade arbetsgruppen en funktion för att matcha förslagen med relevanta passager ur remissvaren. Funktionen gör det också möjligt att söka åt andra hållet, dvs. att utifrån ett utdrag ur ett remissvar söka efter förslag som liknar den delen av remissvaret. Användaren kan filtrera på ifall träffarna ska sökas inom samma utredning som förslaget, bara på andra utredningar eller både inom och utom utredningen.

Sannolikt är det denna matchningstyp som är användbar för flest användare, och det är också den vi har kvalitetstestat mest. Bortom den här semantiska matchningen har vi i projektet för denna matchningstyp även provat att identifiera synpunkter på förslag med hjälp av generativa språkmodeller. Detta beskrivs mer i avsnitt 2.3.

Matchningar mellan förslag och utdrag ur konsekvensbeskrivningar

Om användaren är nyfiken på vilka konsekvenser som diskuteras i samband med ett visst förslag går det att göra en sökning från ett förslag till ett utdrag ur en konsekvensbeskrivning. Sökningen går också att göra åt andra hållet, dvs. att utifrån en viss passage ur ett konsekvensutredningskapitel hitta vilka förslag som ligger semantiskt närmast.

Matchningar mellan remissvar och konsekvensbeskrivningar

Ett syfte med att skicka förslag om nya regler på remiss är att få konsekvenser av förslagen belysta. Det finns därför en stark koppling mellan de konsekvensutredningar som finns i statliga utredningar och remissvaren. Därför har vi utvecklat möjligheten att söka matchningar mellan delar av ett remissvar och delar ur en konsekvensutredning, och vice versa.

Arbetsgruppen har inte hunnit testa de sista två matchningstyperna med andra användare, men vi har ändå lagt ut dem i prototypen så att alla intresserade kan testa funktionerna.

3.5 Automatisk uppdatering av datakällorna

I den ursprungliga prototypen krävdes manuell uppdatering av data för att nya utredningar och remissvar skulle finnas sökbara i tjänsten. Detta gjordes inte regelbundet, utan när behov eller tid fanns.

Under projektet har arbetsgruppen skapat en funktion som automatiserar detta, så nu sker webbskrapning, databehandling och indexering automatiskt en gång i veckan. En risk med detta tillvägagångssätt är att det blir problem i processen om Regeringskansliet förändrar strukturen på sin webbsida.

Den automatiska uppdateringen, som alltså delvis sker från en källa som inte är byggd för strukturerad och löpande delning av information, innebär därför en risk för t.ex. felaktiga remissinstanser och döda länkar eftersom vi inte gör manuella kvalitetskontroller. Data om utredningarna hämtar vi via ett API till riksdagens öppna data, vilket är mindre sårbart.

3.6 Länkar till sparade sökningar

Genom att kunna spara en sökning i form av en länk kan man enklare guida användare till vissa specifika användningsområden, och lättare sprida dessa. Det går att hitta många sådana möjliga exempel. Ett är att en myndighet på sitt intranät eller webbplats kan ha en direktlänk som visar myndighetens senaste remissvar. Ett annat exempel skulle kunna vara en sökning på vissa nyckelord, t.ex. klimat*, miljö* och natur*, som myndigheten kan länka till på sitt intranät för att medarbetarna ska kunna se de senaste förslagen med koppling till nyckelorden. Länkarna leder då till ständigt aktuella listor, givet att den automatiska uppdateringen fungerar (se avsnitt 3.5).

Dessa exempel och många fler är möjliga i den nya prototypen eftersom alla sökningar genererar en unik URL i webbläsarens adressfält, som man enkelt kan kopiera och använda i en länk.

3.7 Övrig vidareutveckling av användargränssnittet

Anpassningen av gränssnittet utifrån de förändringar som vi har gått igenom i kapitel 3 har bl.a. inneburit att arbetsgruppen skapat olika flikar för de olika datatyperna. Utöver denna anpassning till ny funktionalitet och data har vi också vidareutvecklat användargränssnittet på bland annat följande sätt:

- Det är nu möjligt att gruppera sökträffarna per utredning och ”fälla in” de aktuella förslagen ur listan.
- De senaste förslagen och utredningarna presenteras direkt redan på startsidan, och sökningen blir därför mer av en filtrering.
- Vi har lagt till närmast använda rubrik ovanför förslagen. Detta ger ofta en kort sammanfattning av den text som finns i förslagsrutan.

3.8 Ytterligare funktioner för att förbättra Hitta remissvar

Arbetsgruppen har inte kunnat genomföra alla förslag på funktioner som har framkommit utan vi har försökt prioritera de viktigaste eller enklaste. I det här avsnittet beskriver vi förslag på ytterligare funktioner som vi hoppas kommer kunna genomföras vid en eventuell ytterligare vidareutveckling av Hitta remissvar.

3.8.1 Koppla fler datakällor till applikationen

Enligt enkäten till användarna är en av de mest önskade vidareutvecklingarna av tjänsten att koppla fler datakällor till applikationen för att möjliggöra sökningar i andra sorters publikationer, exempelvis rapporter från Statskontoret, Riksrevisionen, utvärderingsmyndigheter eller de stora myndigheternas analysavdelningar.

Arbetsgruppen anser att det är en god idé, men att det ligger utanför avgränsningarna för den här prototypen som är inriktad på just remissvar. Intresset för liknande söktjänster för bredare dataunderlag indikerar att vi har bidragit till syftet med projektet, dvs. att visa på potentialen i att använda AI-baserade språkmodeller i statsförvaltningen.

3.8.2 Tillhandahålla API

Arbetsgruppen har fått förfrågningar om att tillhandahålla data från Hitta remissvar via ett s.k. API, för att förenkla för andra som vill utveckla liknande tjänster. Det är något vi ser positivt på men inte haft möjlighet att utveckla under denna projektfas. Se vidare resonemang i kapitel 5 om öppna data och öppen kod.

3.8.3 Inkludera äldre remissvar

En återkommande synpunkt från användare är att det vore värdefullt att även kunna söka på äldre remissvar än från 2016. Det skulle i dagsläget vara ett för stort jobb att få fram dessa eftersom de inte finns publicerade på regeringen.se. Därmed skulle det kräva ett omfattande insamlingsarbete.

3.8.4 Inbyggd funktion för användarfeedback

För att kunna utveckla applikationens träffsäkerhet och användarvänlighet ville arbetsgruppen utveckla en funktion för användarfeedback. Funktionen skulle ge användaren möjlighet att betygsätta hur bra och relevant en matchning är. Funktionen skulle även kunna ge användaren möjlighet att motivera sina synpunkter eller lämna kommentarer. Löpande insamling av feedback skapar dock förväntningar på ett tydligt åtagande och resurser att drifva och vidareutveckla tjänsten (se kapitel 4).

3.8.5 Sortera bort remissvar utan synpunkter

Arbetsgruppen ville göra det möjligt att automatisk sortera bort remissvar där remissinstansen helt avstår från att lämna synpunkter. Även om vi har testat funktionaliteten (se avsnitt 2.3) så har vi inte lyckats driftsätta någon sådan lösning.

3.8.6 Remissammanställning, gruppering och identifiering av mönster i data

En möjlig framtida funktion för Hitta remissvar är att användaren ska kunna få en snabb och förenklad bild av remissopinionen i en fråga med hjälp av sentimentanalys och generativ AI. För en sådan sammanställning av remissvar vore det fördelaktigt att kunna sortera informationen, exempelvis utifrån typ av organisation och sökresultatet per förslag.

En närliggande idé är att datadrivet utforska vilka mönster som finns i data (oberoende av strukturen i betänkandet och remissinstanserna). Ett exempel på ett sådant mönster är vilka frågor eller delar som de flesta remissinstanserna har synpunkter på. Att översiktligt visa detta skulle ge en snabb introduktion för den som vill sätta sig in i en fråga och få en bild av vad som tycks vara de centrala tviste-frågorna i förslagen från en viss utredning.

Andra möjligheter som skulle kunna vara intressanta att utforska vidare är att försöka illustrera vilka remissvar som ligger närmast varandra temamässigt, dvs. har synpunkter på samma frågor. Ett nästa, men sannolikt svårare steg, vore att se hur nära instanserna ligger varandra i sina ståndpunkter inom dessa frågor.

3.8.7 Ytterligare förfining av användargränssnittet

Återkoppling från användare är ett mycket viktigt verktyg för att utveckla användarvänliga gränssnitt. Genom att tillhandahålla Hitta remissvar i ett öppet webbgränssnitt möjliggör vi för alla intresserade att testa tjänsten. Det innebär att det finns stora möjligheter att samla värdefulla användarinsikter för fortsatt förbättring av användargränssnittet.

4 Att gå från prototyp till färdig tjänst

ESV och Tillväxtverket har under detta projekt vidareutvecklat en prototyp för att visa på en rad nyttor för myndigheter, organisationer och samhället i stort. Syftet med projektet har inte varit att ta fram en färdigutvecklad tjänst för permanent användning. Detta beror delvis på att det inte ingår i uppdraget för ESV:s datalabb där applikationen har utvecklats. Datalabbets fokus är att visa på möjligheter genom just prototyper. Det ingår däremot i uppdraget att beskriva vilka steg som krävs för att ta prototypen vidare till en färdig tjänst, vilket vi redogör för i detta kapitel.

4.1 Ytterligare anpassningar krävs

Att ta fram en prototyp är ofta ett första steg i att genomföra ett utvecklingsprojekt. Men att gå från en prototyp till en färdig tjänst innebär mer än att bara göra ”klart” ett system. Det handlar om att säkerställa att tjänsten fungerar smidigt, uppfyller såväl tekniska som juridiska krav och kontinuerligt kan anpassas till förändringar i omvärlden. Det innefattar flera viktiga steg och överväganden och är en process som kräver tid, resurser och noggrann planering för att säkerställa en lyckad implementering och driftsättning av tjänsten. Ett första steg i en sådan process är att hitta en aktör som är villig att ta sig an detta.

En prototyp är i huvudsak en första version som fokuserar på grundläggande funktioner och kan exkludera flera viktiga aspekter, t.ex. sekretesshantering eller tillgänglighetskrav. För Hitta remissvar är all data offentliga handlingar utan problem med sekretess, men arbetsgruppen har ändå behövt förhålla sig till frågan om personuppgifter i remissvaren (se avsnitt 3.2.1).

Vi har också försökt hålla oss till tillgänglighetskraven vid utvecklingen, men utan att göra någon djupare utvärdering av hur väl vi lyckats. Med en färdig tjänst behöver man säkerställa stabilitet, datatillgänglighet och följsamhet till tillgänglighetskraven, samt göra ytterligare anpassningar till förvaltningsmässiga och juridiska krav. Tjänsten måste även vara anpassad till den förvaltande organisationens digitala ekosystem. Det krävs kompetens och resurser för att förvalta och vidareutveckla tjänsten och säkerställa integrationer till andra system.

Förvaltning av en tjänst innebär att ha ansvar för dess kontinuerliga utveckling och underhåll. Det inkluderar att hantera buggar, anpassa tjänsten till förändringar i omvärlden och säkerställa att juridiska krav uppfylls. Det är en kontinuerlig process som kräver samarbete mellan it och verksamhet, tydlig ansvarsfördelning och överenskommelser om allt från backuprutiner till supporttider.

Under övergången från prototyp till färdig tjänst krävs också en granskning av tekniska aspekter, såsom att säkerställa systemets godkännande enligt organisationens utvecklingsriktlinjer och eventuella behov av att anskaffa nya licenser eller infrastruktur.

I det specifika fallet med Hitta remissvar, där prototypen bygger på webbramverket Angular och sökmotorn Elasticsearch, handlar övergången till en färdig tjänst främst om att säkra dataflödet och eventuella behov av ny it-infrastruktur. Det kräver också övervakning av datainsamling och kanske anskaffning av nya driftsmiljöer, molntjänster eller licenser.

4.2 Kostnader för att permanenta tjänsten

Övergången från prototyp till en färdig tjänst innebär såväl initiala utvecklingskostnader för att färdigställa och anpassa tjänsten, som löpande kostnader för att driva och förvalta den. Engångskostnader för infrastrukturinköp kan också tillkomma.

4.2.1 Utvecklingskostnader

Att ta prototypen Hitta remissvar till en färdig tjänst kräver betydande utvecklingsinsatser. Beräkningen av utvecklingskostnaderna är komplex och involvera många faktorer. En mycket grov uppskattning pekar på en ungefärlig utvecklingskostnad på 500 timmar. Detta inkluderar anpassning av programmet, byggande av ny infrastruktur samt att säkra överensstämmelse med organisationens anskaffningsprocess och att systemet fungerar väl i produktionsmiljön.

4.2.2 Löpande förvaltning och driftskostnader

När tjänsten är igång krävs löpande förvaltning och drift och någon som är utpekat ansvarig för tjänsten. För att hålla systemet igång och säkerställa kontinuerlig prestanda behövs det regelbunden övervakning, buggfixar, uppdateringar och säkerhetsåtgärder. Det kan t.ex. handla om att någon ingående komponent (hypotetiskt exempel i detta fall den BERT-modell som används) slutar underhållas och inte längre går att få support på.

Kostnader för löpande förvaltning och drift inkluderar också att hantera förväntade och oväntade avvikelser i data. Som en mycket grov uppskattning kan det för Hitta remissvar kräva omkring 100 timmar per år för löpande drift och underhåll.

4.2.3 Infrastrukturkostnader

Vid behov av ny infrastruktur, exempelvis en Linux-maskin för att stödja tjänsten, tillkommer kostnader för inköp och eventuell installation. Dessa kostnader varierar beroende på specifikationer och vilka behov som måste tillgodoses. Det är också viktigt att beakta eventuella kostnader för licenser och support för mjukvara eller

verktyg som kan krävas för den nya infrastrukturen. Molntjänster kan i det här fallet vara ett alternativ till att använda egen infrastruktur.

5 Varför öppna data och öppen källkod ännu inte är fullt implementerade

Syftet med projektet Hitta remissvar är att utforska potentialen av att göra befintliga uppgifter lättillgängliga genom att förenkla sökning och koppling av data. I projektet har arbetsgruppen delvis kunnat uppfylla de efterfrågade funktionerna som framkommit i dialog med olika intressenter. Denna process har understrukit vikten av att dela kunskap och möjliggöra för andra att utveckla ytterligare användningsområden för tekniken och datamängden.

5.1 Öppna data

Genom projektets gång har ESV:s datalabb fått förfrågningar från forskare och myndigheter om att få tillgång till de data som ligger till grund för Hitta remissvar. Dessa förfrågningar pekar på vikten av att publicera så mycket data som möjligt som öppna data, som är tillgängliga för alla att använda och återanvända. Det kan bl.a. bidra till ny kunskap och innovation. ESV:s digitala policy stödjer denna ambition, i linje med regeringens uppmaningar om att använda data som en resurs för innovation och användarcentrerad service.¹⁷

Även om ESV regelbundet publicerar ett urval av data som öppna data¹⁸, är det en utmaning att göra detta med all data som finns tillgänglig hos myndigheten. Det beror delvis på att vissa data inte är kvalitetssäkrad och heller inte uppdateras kontinuerligt. Därför kan ESV behöva säkerställa informationsansvar och informationssäkerhetsklassificering innan publicering.

För Hitta remissvars data finns specifika utmaningar. Till exempel måste vi säkerställa att ESV har rätt att publicera data från externa källor och hantera eventuella risker som detta kan medföra. Dessutom behöver vi säkerställa att hanteringen av personuppgifter är i enlighet med dataskyddsförordningen. Vi arbetar för att adressera dessa utmaningar och har som mål att publicera data från Hitta remissvar som öppna data under 2024.

5.2 Öppen källkod

Hitta remissvar utvecklas av Ekonomistyrningsverkets datalabb, en avdelningsöverskridande funktion som fokuserar på utveckling i syfte att förstå potentialen med AI i ESV:s verksamheter eller statsförvaltningen i stort, och har inte i uppdrag att förvalta permanenta tjänster. Öppen källkod är ett värdefullt verktyg för att

¹⁷ ESV (2021), *Digital policy*. Dnr F-2021-11.

¹⁸ ESV, *Ekonomistyrningsverkets öppna data*, <https://www.esv.se/statistik-och-data/oppna-data/>. Hämtat 20 februari 2024.

säkerställa att tjänsten kan leva vidare, även om projektet avslutas. ESV har redan fått förfrågningar om att göra koden öppen, och vi arbetar för att göra detta möjligt.

Flera myndigheter, inklusive Myndigheten för digital förvaltning (Digg), eftersträvar att utveckla och använda öppen källkod där det är möjligt.¹⁹ Även om det inte finns en lagstiftning som kräver detta, är det ett värdefullt sätt för den offentliga sektorn att dela resurser och kunskap och att öppna upp fler innovationsmöjligheter. ESV:s kärnuppdrag inkluderar it-utveckling, och ESV bedömer att koden för Hitta remissvar och andra projekt kan publiceras utan säkerhetsrisker.

¹⁹ Myndigheten för digital förvaltning, Digg (2022), [Riktlinjer för utveckling och publicering av öppen programvara](#).

6 Lärdomar och rekommendationer för användning av textdata och AI-verktyg i offentlig sektor

I detta kapitel delar vi med oss av lärdomar som arbetsgruppen tror kan vara till nytta för andra som arbetar med liknande utvecklingsprojekt.

6.1 Databeredskapen måste bli bättre för att ta tillvara på potentialen i offentlig text

ESV har under flera år pekat på att databeredskapen i myndigheter behöver förbättras och även i detta projekt gör vi samma konstaterande.²⁰ Det är viktigt att myndigheter ser text som data, tillgängliggör befintliga texter i maskinläsbara format och skapar rutiner för att säkerställa tillgängligheten till data framöver.

6.1.1 Några av nuvarande problem

En av de utmaningar som har identifierats i arbetet med Hitta remissvar är processen att tillgängliggöra data från pdf-filer, som ofta publiceras på offentliga webbplatser. För att bearbeta dessa data krävs webbskrapning och processering, vilket är mer krävande än att använda data i textformat via API. Webbskrapning är också en sårbar metod när strukturen på den webbplats som skrapas ändras.

Andra exempel på dataproblem som arbetsgruppen stött på under projektets gång och som illustrerar behovet av förbättringar och en systematisk hantering av frågorna:

- **Vissa remissinstanser publicerar otillgängliga dokument**, t.ex. inskannade fullmäktigeprotokoll som är svåra att läsa maskinellt. Den första utmaningen är därför att kunna extrahera uppgifterna från dokumenten där de är inlåsta. Under de senaste åren har nya modeller utvecklats för att automatisera extraheringen från ostrukturerade dokument såsom GPT-4-vision²¹ eller donut²². Dessa modeller har framgångsrikt testats i andra projekt i ESV:s datalabb. Men även med sådana verktyg krävs betydande ansträngningar för att tillgängliggöra data.
- **Bristen på standardisering** i benämning och filhantering hos remissinstanserna. Felstavningar, namnbyten och inkonsekvent användning av benämningar och förkortningar kan leda till svårigheter i att hitta alla relevanta dokument.

²⁰ Se till exempel ESV, *Vi behöver öka databeredskapen i staten*. <https://www.esv.se/press-och-publicerat/nyhetsbrev/ledare-november-2023/> Uppdaterad 2023-11-30. Hämtad 13 februari 2024. Och ESV, *Att bli en datadriven myndighet*. <https://forum.esv.se/verksamhetsutveckling/digitalisering-och-ai/Att-bli-en-datadriven-myndighet/> Uppdaterad 2023-02-23. Hämtad 13 februari 2024.

²¹ OpenAI, *GPT-4-vision*. Hämtad 13 februari 2024.

²² Arxiv.org, *OCR-free Document Understanding Transformer*, <https://arxiv.org/abs/2111.15664>. Uppdaterad 6 oktober 2022. Hämtad 6 februari 2024.

- **Manuella processer i uppladdning** av dokument på offentliga webbplatser kan leda till fel. Exempelvis kan en felaktig rubrik eller fil orsaka problem i datahanteringen, vilket vi bl.a. observerade när en utredning felaktigt laddades upp under en annan rubrik och genererade inkonsekventa data.

6.1.2 Budskap för bättre databeredskap

Vårt projekt delar de slutsatser om databeredskap som framförs i skriften ”*Databeredskap för språkteknologiska tillämpningar - Fallstudier från offentlig sektor*”²³ som tagits fram inom ramen för AI Swedens projekt Databeredskapsverkstaden där bl.a. ESV medverkade. För att öka kännedomen om och spridningen av dessa budskap återger vi dem, i urval, här.

Text är data

Det behövs en ökad förståelse inom offentlig sektor att data inte är synonymt med sifferdata utan att även text är data, och att textdata inte är samma sak som textdokument. Textdata behöver extraheras från dokument, som kan ha olika grad av maskinläsbarhet beroende på filformat och olika grad av användbarhet beroende på vilken metadata som texten är taggad med.

Textdokument i PDF-format har låg grad av maskinläsbarhet.

Textdata behöver extraheras från dokumentet och PDF är inte ett format som underlättar detta. Samtidigt så är PDF ett av de absolut vanligaste formaten för publicering av textdokument i offentlig sektor. När PDF ändå används för att lagra textdata bör det vara i tillgänglighetsanpassat format (PDF/UA) med metataggar som tydligt indikerar textens struktur och uppdelning.

Vi behöver låsa upp gammal data ur befintliga system och framtidssäkra ny data.

Att utveckla en kapacitet att hantera befintliga datakällor och extrahera och behandla textdata från dem är ett första steg i god databeredskap. I ett längre perspektiv är det viktigt att också se över och utveckla de processer där ny textdata skapas och lagras så att framtida data är av bättre kvalitet, mer heltäckande, mer tillgänglig, mer maskinläsbar och har mer användbar metadata.

²³ Databeredskapsverkstaden (2023), [Databeredskap för språkteknologiska tillämpningar - Fallstudier från offentlig sektor](#).

NLP-tillämpningar²⁴ kan inte bli bättre än den data de bygger på.

Avancerade metoder kan bara delvis och i vissa fall kompensera för brister i datakvalitet. Med rätt data i tillräcklig kvalitet kan samma problem ofta lösas på betydligt enklare sätt samtidigt som det möjliggör mer avancerade tillämpningar.

Samarbete över myndighetsgränser behövs.

Språkteknologiska tillämpningsmöjligheter och intressanta textdatakällor är ofta inte tydligt kopplade till endast en offentlig aktör. Det har därför hänt, och kommer fortsätta att hända, att olika aktörer bedriver snarlika projekt utan att vara medvetna om varandra. Detta innebär ett ineffektivt dubbelarbete och en missad möjlighet till samverkan. Det behövs därför bättre kommunikation inom det offentliga kring vilka språkteknologiska projekt som drivs och planeras samt bättre styrning och ökade medel för gemensamma projekt.

6.1.3 Åtgärder som enskilda remissinstanser kan göra för att bidra till bättre databeredskap och funktionalitet i Hitta remissvar

Information om vad myndigheter själva kan göra för att bidra till bättre databeredskap och tillgänglighet finns bl.a. i Digg:s vägledning för öppna data²⁵, i EU:s tillgänglighetsdirektiv²⁶, på dataportalen²⁷, i wcag-kraven²⁸ och i eSams rapporter²⁹. Digital tillgänglighet har i stor utsträckning med maskinläsbarhet att göra, och går därför ofta hand i hand med databeredskapsfrågor.

Specifikt för Hitta remissvar så blir sökresultaten i prototypen bättre om remissinstanserna gör följande saker:

- Publicera remissvaren i ett tillgänglighetsanpassat format med rätt taggning av rubriker och annan metadata.
- Ange tydligt vilket förslag som synpunkten hör till, gärna med den exakta lydelsen av förslaget och genom att återge samma kapitelrubrik som i betänkandet. Detta framgår redan i regeringens PM om att svara på remiss.³⁰

²⁴ NLP är en förkortning för Natural Language Processing. På svenska pratar man om språkbaserade AI-modeller eller språkteknologi som handlar om hur datorprogram kan analysera, producera, modifiera och svara på mänskligt text och mänskligt tal.

²⁵ Myndigheten för digital förvaltning, Digg, Öppna och delade data, <https://www.digg.se/kunskap-och-stod/oppna-och-delade-data>. Uppdaterad 14 augusti 2023. Hämtad 6 februari 2024.

²⁶ Europaparlamentets och rådets direktiv (EU) 2019/882 av den 17 april 2019 om tillgänglighetskrav för produkter och tjänster. <https://eur-lex.europa.eu/legal-content/SV/ALL/?uri=CELEX:32019L0882>

²⁷ Myndigheten för digital förvaltning, Digg, Sveriges dataportal, <https://www.dataportal.se/>. Hämtad 6 februari 2024.

²⁸ Myndigheten för digital förvaltning, Digg, Det här är EN 301 549 och WCAG, <https://www.digg.se/webbriktlinjer/lagar-och-krav/det-har-ar-en-301-549-och-wcag>. Uppdaterad 15 december 2023. Hämtad 6 februari 2024.

²⁹ T.ex. eSam (2021), [Rapport om nuläget Tillgängliggörande och vidareutnyttjande av data](#).

³⁰ "Som rubriker i remissvaret passar det ofta att använda rubrikerna till avsnitten i förslaget. Om avsnitten i förslaget är numrerade kan det hänvisas till samma numrering i remissvarets rubriker". Hämtat från SB PM 2021:1, Statsrådsberedningen, *Svara på remiss*. Regeringskansliet.

- Gör inte rena sammanfattningar av SOU:n i remissvaret, särskilt inte av delar som myndigheten inte har någon uppfattning om (då det gör att man får träffar i tjänsten även på sådant som remissvaret inte har någon synpunkt på).

6.2 Stort omvärldsberoende och hastigt ändrade förutsättningar

I projektet har vi med Hitta Remissvar som exempel illustrerat hur språkbaserade AI-modeller kan användas för att behandla offentliga texter som data och förhoppningsvis effektivisera analysen av texterna. Genom att testa avancerade språkmodeller som GPT-4 och GPT-SW3 har projektet undersökt möjliga tillämpningar som kan komma att bli betydligt vanligare och mer välutvecklade inom en snar framtid.

Möjligheterna att dra nytta av ny teknik är starkt kopplade till ett omfattande nätverk av öppna resurser³¹ och den kontinuerliga utvecklingen inom AI-communityn, inklusive bidrag från organisationer som Kungliga Bibliotekets KB-labb, AI Sweden, eSam³², Digg och andra.

En stor del av tekniken utvecklas i USA eller av stora internationella samarbeten, vilket tydliggör projektets starka koppling till globala innovationstrender. Utveckling och användning av AI medför beroenden, etiska frågor och risker som kan vara nya för myndigheter att hantera. Hur påverkar valen av modeller, data och utvärdering vilka resultat AI:n visar? Exempelen på diskriminerande och på andra sätt bristande AI-tillämpningar är många. Hur kan vi förstå, välja, använda och förklara tekniken på ett lämpligt sätt? Det ställer krav på kompetensutveckling, transparens och involvering av flera yrkesgrupper och en lyhördhet för samhällsdiskussionen om AI.

Vi kan förvänta oss att regleringen av AI kommer att förändras betydligt de kommande åren, med pågående förhandlingar på EU-nivå om en AI-förordning som kommer att fastställa regler och standarder för användningen av AI. Detta kommer att påverka hur offentliga förvaltningar och projekt som vårt kan nyttja AI-teknik. Därför krävs kontinuerlig anpassning av våra metoder och verktyg för att följa dessa regler. Den dynamiken utgör en utmaning med tanke på den snabba utvecklingen både inom AI-tekniken och dess reglering. Det är avgörande att kontinuerligt följa och anpassa sig till omvärlden för att säkerställa att vår användning av AI är ändamålsenlig och följer gällande regler.

³¹ T.ex. verktyg som pytorch, pandas och opencv.

³² Under verksamhetsåret 2023/2024 driver eSam sju aktiviteter relaterade till AI, bland annat framtagande av en myndighetsgemensam processbeskrivning för utveckling av AI och analyser av utvecklingen av generativ AI och möjlig påverkan på myndigheterna. Medlemmarna i eSam samverkar löpande med trendanalys och omvärldsbevakning och publicerar en så kallad radarbild med internationella och nationella initiativ som påverkar myndigheternas digitaliseringsarbete.

6.3 Även ofärdiga och ickeperfekta prototyper kan ge konkret nytta

Att ta fram något konkret kan, även om det inte är perfekt, skapa direkt nytta och göra att användarna kan skapa sig en bild av de möjligheter som användning av textdata och språkbaserade AI-modeller kan ge. En faktisk tjänst, om än i prototypform är ett illustrativt sätt att visa nyttor. Vi ser att prototyper kan vara ett utmärkt komplement till utredningsverksamhet. Även små steg i innovationsprocesser tror vi leder till förbättringar som ytterst gagnar medborgare och samhället i stort.

Att prototypen har hundratals besök varje månad, får positiv respons vid demonstrationer och att vissa myndigheter uppdaterat sin remissvarsrutin med hänvisning till Hitta remissvar är några indikationer på att tjänsten ger nytta.

En annan nytta som utvecklingsarbetet redan har gett upphov till är att den datamängd som arbetsgruppen tagit fram genom webbskrapning och processering av remissvaren kan återanvändas i ett forskningsprojekt från ESO med fokus på remissförfarandet.³³ Detta illustrerar värdet av att göra datamängder maskinläsbara. När data finns i maskinläsbart och strukturerat format skapas nya möjligheter till nytta, innovation och forskning som den ursprungliga producenten kanske inte hade föreställt sig.

Hitta remissvar är bara ett litet exempel på en tjänst man kan bygga med en bråkdel av den text som offentlig sektor producerar. Ambitionen med projektet har varit att dels inspirera fler till att bygga nya tjänster för andra behov, dels bidra till att fler myndigheter förbättrar hur de tillhandahåller textdata så att mer innovation blir möjlig.

6.4 Varken människa eller maskin kan dra tydliga slutsatser av vaga remissvar

En av frågorna för ovan nämnda ESO-projekt med fokus på remissförfarandet är hur pass tydliga ställningstaganden som remissinstanserna ger. Det är en fråga vi också närmat oss i projektet när vi lät både utredare och språkmodeller bedöma huruvida remissvaren förhöll sig positiva, negativa eller neutrala till ett visst förslag. En av slutsatserna var att det ibland är svårt att bedöma för såväl människa som maskin.

Detta visar på att det inte räcker att göra remissvar maskinläsbara för att enkelt sammanställa remissinstansernas inställning till ett förslag. De som skriver remissvaren kan också behöva bli tydligare, både i hur de uttrycker synpunkter och hur det tydliggörs huruvida synpunkterna gäller samtliga förslag eller delar av

³³ Projektet nämns bland nya ESO-projekt i [ESO:s nyhetsbrev från november 2023](https://eso.expertgrupp.se/nyhetsbrev-fran-eso-25/).
<https://eso.expertgrupp.se/nyhetsbrev-fran-eso-25/>. Publicerad 2 november 2023. Hämtad 15 februari 2024.

förslagen. Där kan både tydligare argumentation, ökad användning av metataggar och/eller olika digitala verktyg vara vägar framåt.

ESV:s datalabb har experimenterat med en GPT som automatiskt bedömer remissvar utifrån Regeringskansliets riktlinjer om att svara på remiss. Tjänsten ger bl.a. en indikation på ställningstagandets tydlighet. Ett sådant verktyg skulle kunna vara till stöd för dem som skriver remissvar.

Referenser

Arxiv.org, *OCR-free Document Understanding Transformer*,
<https://arxiv.org/abs/2111.15664>. Uppdaterad 6 oktober 2022. Hämtad 6 februari 2024.

Databeredskapsverkstaden (2023), *Databeredskap för språkteknologiska tillämpningar - Fallstudier från offentlig sektor*

eSam (2021) *Rapport om nuläget Tillgängliggörande och vidareutnyttjande av data*

ESO, Nyhetsbrev från ESO. <https://eso.expertgrupp.se/nyhetsbrev-fran-eso-25/>.
 Publicerad 2 november 2023. Hämtad 15 februari 2024.

ESV (2021), *Digital policy*. Dnr F-2021-11.

ESV, *Att bli en datadriven myndighet*.
<https://forum.esv.se/verksamhetsutveckling/digitalisering-och-ai/Att-bli-en-datadriven-myndighet/> Uppdaterad 2023-02-23. Hämtad 13 februari 2024.

ESV, *Vi behöver öka databeredskapen i staten*. <https://www.esv.se/press-och-publicerat/nyhetsbrev/ledare-november-2023/> Uppdaterad 2023-11-30. Hämtad 13 februari 2024.

ESV, *Ekonomistyrningsverkets öppna data*, <https://www.esv.se/statistik-och-data/oppna-data/>. Hämtat 20 februari 2024.

[Europaparlamentets och rådets direktiv \(EU\) 2019/882](#) av den 17 april 2019 om tillgänglighetskrav för produkter och tjänster.

Explain xkcd, 2347: *Dependency*,
https://www.explainxkcd.com/wiki/index.php/2347:_Dependency. Uppdaterad 12 oktober 2023. Hämtad 7 februari 2024.

Förordningen (2009:145) med instruktion för Tillväxtverket.

Förordningen (2016:1023) med instruktion för Ekonomistyrningsverket.

IDG, *ordlista över ord och uttryck i IT-branschen*. Hämtad 30 januari 2024.

Internetstiftelsen, *Var fjärde internetanvändare väljer bort cookies* (Diagram 2.7 Klickar ur cookies – fördelat på kön/generation),
<https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2022/vad->

[begransar-var-internetanvandning/var-fjarde-internetanvandare-valjer-bort-cookies/#vilka-klickar-ur-cookies](#). Hämtad 8 februari 2024.

KB-labb på HuggingFace, *Swedish BERT Models*, <https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/amnesovergripande-statistik/namnstatistik/>. Hämtad 30 januari 2024.

Malmsten, M., Börjeson, L., och Haffenden, C. (2020), *Playing with Words at the National Library of Sweden--Making a Swedish BERT*.

Myndigheten för digital förvaltning, Digg (2022) *Riktlinjer för utveckling och publicering av öppen programvara*.

Myndigheten för digital förvaltning, Digg, *Det här är EN 301 549 och WCAG*, <https://www.digg.se/webbriktlinjer/lagar-och-krav/det-har-ar-en-301-549-och-wcag>. Uppdaterad 15 december 2023. Hämtad 6 februari 2024.

Myndigheten för digital förvaltning, Digg, *Sveriges dataportal*, <https://www.dataportal.se/>. Hämtad 6 februari 2024.

Myndigheten för digital förvaltning, DIGG, *Öppna och delade data*, <https://www.digg.se/kunskap-och-stod/oppna-och-delade-data>. Uppdaterad 14 augusti 2023. Hämtad 6 februari 2024.

OpenAI, *GPT-4-vision*. <https://platform.openai.com/docs/guides/vision>. Hämtad 13 februari 2024.

Policy Impact (2022), *De samhällsekonomiska effekterna kopplat till Kungliga bibliotekets AI-baserade språkmodeller*

Riksdagens öppna data, *Användningsvillkor*. Hämtad 30 januari 2024.

SB PM 2021:1, Regeringskansliet, Statsrådsberedningen, *Svara på remiss – om remisser av betänkanden och andra förslag från Regeringskansliet*

Statistiska centralbyrån, *Namnstatistik*, <https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/amnesovergripande-statistik/namnstatistik/>. Hämtad 30 januari 2024.

Vinnova, *GPT-SW3: en svensk basmodell för texthantering*, <https://www.vinnova.se/p/gpt-sw3-en-svensk-basmodell-for-texthantering/>. Uppdaterad 30 september 2022. Hämtad 26 februari 2024.

Bilaga 1 – Möjliggörande aktörer

Hitta remissvar hade inte kunnat utvecklas utan de komponenter som tillhandahålls öppet av KB, Kungliga Biblioteket, (språkmodellen KB-BERT) och Riksdagens öppna data (textdata från utredningarna). Inom ramen för projektet har arbetsgruppen träffat dessa möjliggörande aktörer för att uttrycka vår uppskattning för deras arbete, visa upp vad vi kunnat göra tack vare dem samt inhämta synpunkter och utbyta erfarenheter. Nedan redogör vi mer i detalj för dessa aktörers bidrag och arbete.

Kungliga Biblioteket

Arbetsgruppen träffade Kungliga Bibliotekets datalabb den 23 november 2023 för att presentera tjänsten, utbyta erfarenheter och identifiera samarbetsmöjligheter.

Kungliga bibliotekets datalabb, KB-labb, är en nationell infrastruktur för digital forskning och utveckling inom artificiell intelligens. Grunden för KB-labb är formuleringen i KB:s instruktion om att KB ska bidra till svensk forskningskvalitet. Målgruppen för KB-labb är akademien, och den första uppgiften KB-labb löste var att ge storskalig digital åtkomst till KB:s samlingar. Det man idag är mest känd för är nog dock utvecklingen och tillgängliggörandet av högkvalitativa AI-baserade språkmodeller på svenska. Modellerna bygger på Kungliga Bibliotekets unika och stora samlingar, som KB-labb är den enda aktör som har lovlig tillgång att träna på. Utvecklingen är dessutom väldigt kostnadseffektiv eftersom så mycket data redan finns. Språkmodellerna är öppet publicerade för allmänheten och gratis att använda.

KB-BERT har laddats ner ca 3 miljoner gånger, och används av allt ifrån stora myndigheter till små inkassoföretag. Det finns dock ingen samlad information om vilka som använder KB-labb:s modeller, utan den kännedom som finns baseras på vad som inkommit anekdotiskt.

Stor samhällsnytta av KB:s modeller

Att Hitta remissvar bygger på KB-BERT är ett litet exempel på all den nytta som KB-labbs AI-modeller leder till. I rapporten De samhällsekonomiska effekterna kopplat till Kungliga bibliotekets AI-baserade språkmodeller från 2022 gör konsultbolaget Policy Impact en uppskattning om det aggregerade framtida värdet av ifall stora statliga myndigheter skulle använda sig av KB:s modeller.

Utifrån ett realistiskt framtidsscenario där större statliga myndigheter med fler än 500 årsarbetskrafter - 46 stycken totalt - successivt låter sin kunskapsintensiva arbetskraft nyttja AI-baserade språkmodeller för textanalys och sökning beräknas detta kunna generera en sammantagen samhällsekonomisk nettomarginalnytta på närmare 4,2 miljarder kronor över

en femårsperiod mätt i dagens penningvärde. Detta innefattar en halvering av tiden som läggs på textsökning och analys - från 20 till 10 procent av arbetstiden - samt bidrar till en mindre reduktion av externa avgifter vilket gagnar näringslivet och allmänheten. Utifrån ett realistiskt antagande om att KB-labbs språkmodeller utgör basen för en femtedel av tillämpningarna i framtidsscenarioet skulle närmare 840 miljoner kronor av den totala nettomarginalnyttan kunna tilläggas till KB-labb. Detta skulle utgöra ett extremt positivt nytto-kostnads-ratio för KB-labb givet den låga löpande marginalkostnaden för förvaltning och utveckling av nya modeller.³⁴

Rapporten tar också upp att det skulle krävas mer och bättre samverkan samt delande av kompetens, data och resurser mellan olika aktörer inom offentlig sektor för att kunna bidra till snabbare nyttorealiserings för lägre kostnad.

Arbetsgruppen för Hitta remissvar konstaterar att även mindre myndigheter har analys- och text-intensiva uppgifter som skulle kunna dra nytta av AI-modeller, och att den verkliga potentialen och nyttan sannolikt ligger betydligt högre än de modesta uppskattningar som gjorts i konsultrapporten.

Språkmodellens bidrag till likriktning vs. Representation

En skillnad mellan KB-labbs modeller och de stora modellerna framtagna av internationella företag är att KB-labb inte censurerar potentiellt problematiskt innehåll. KB-labbs hållning är att fokus bör ligga på att modellerna ska vara representativa snarare än att konformera sig till normer och förväntningar och på så vis riskera att bidra till likriktning och strömlinjeformning av språk och i förlängningen tankar.

Ett historiskt exempel på denna problematik är att om Fredrika Bremer hade blivit censurerad i sin tid, hade innehåll om kvinnors rättigheter kanske ansetts vara provocerande, medan rasismen accepterats som sunt förnuft. Det vi väljer att censurera i dag kommer att skapa historiska avtryck och formas av det samhälle vi lever i.

Bibliotekens uppgift sträcker sig bortom att bara samla in de mest populära böckerna och tidningarna. Det handlar om att vara skicklig på representation, att säkerställa att olika perspektiv och röster ges plats. AI-modeller spelar en betydande roll i detta sammanhang. Utan aktiv utveckling för att inkludera olika perspektiv riskerar de att förstärka och cementera existerande samhällsnormer och åsikter.

³⁴ Policy Impact (2022), [De samhällsekonomiska effekterna kopplat till Kungliga bibliotekets AI-baserade språkmodeller](#).

Riksdagens öppna data

Arbetsgruppen träffade delar av Riksdagsförvaltningens öppna data-team den 15 september. Riksdagens öppna data bygger på den tidigare tjänsten Rixlex, som var ett samarbete mellan Riksdagens förvaltningskontor och Regeringskansliet, i samråd med bl.a. Domstolsverket.

Från början (1993) var det en betaltjänst, som sedan blev gratis (1995) och som senare ersattes av webbplatsen riksdagen.se. År 2010 vidareutvecklades detta till att bli riksdagens öppna data. Det fanns inget formellt beslut eller uppdrag att utveckla riksdagens öppna data, utan initiativet kom inifrån organisationen.

Riksdagens öppna data vet i dagsläget inte vilka som använder sig av deras tjänster, men de funderar på att erbjuda möjligheten att registrera sig som användare så att de kan kommunicera ut eventuella förändringar samt att användarna ska ha en bra ingång för att rapportera buggar och liknande.

Teamet bakom Riksdagens öppna data gör ett manuellt jobb för varje SOU och DS som publiceras. De prenumererar på ett nyhetsbrev från regeringen, laddar hem dokumenten och lägger dem i en katalog. Sedan körs ett jobb som automatiskt konverterar dem till html-format och lägger till metadata (t.ex. titel).

Riksdagens öppna data har inte resurser till egen kontroll av dokumenten, och uppskattar därför felanmälningar. Ibland beror felet på att dokumenten från Regeringskansliet är korrupta och därför inte kan konverteras till html. Trots påpekanden till Regeringskansliet har detta inte korrigerats. Detta hände dock framför allt tidigare.

Regeringskansliet har minskat antalet utredningar inom SOU och DS, och gått över till att göra korta PM istället. Dessa PM blir pdf:er på remissidan på regeringen.se, men i dagsläget hämtas inte dessa hem och tillgängliggörs via Riksdagens öppna data.

Riksdagens Öppna data bedriver ett arbete med att göra flera och modernare API:er. Dessutom pågår ett arbete med att tillgängliggöra och kvalitetsförbättra äldre material. Ambitionen är att tillgängliggöra fullständigt riksdagstryck sedan 1867 i ett första steg. Långsiktigt är förhoppningen att kunna tillgängliggöra även stånds- riksdagens dokument.

Bilaga 2 – Användarfeedbacken i sin helhet

Introduktion

Syftet med projektet Hitta remissvar med AI är att visa på potentialen i att använda AI-baserade språkmodeller i statsförvaltningen. Effektmålet på längre sikt är att utveckla AI-baserade analys- och beslutsmetoder som bidrar till bättre beslutsunderlag och en effektivisering av remisshanteringen i staten.

En viktig del av utvecklingsarbetet har därför handlat om att löpande testa och utvärdera prototypen Hitta remissvar och möjlig framtida funktionalitet. I denna bilaga beskriver vi återkopplingen vi fått från användare av Hitta remissvar.

Om applikationen

Hitta remissvar är en applikation (app) som ger stöd i att hitta och jämföra liknande förslag från olika utredningar samt tillhörande remissvar (se figur 1).

Applikationen underlättar för avancerade sökningar i statliga utredningar som publiceras i departementsserien (Ds) och serien för statens offentliga utredningar (SOU). Målsättningen är att göra utredningar som remitteras av regeringen mer överskådliga och tillgängliga för analys. Applikationen har bl.a. en funktion som automatiskt identifierar och extraherar de förslag som presenteras i utredningarna.

I applikationen görs matchningar mellan förslagen i olika utredningar. Det gör att användaren kan hitta tidigare remissvar som matchar nya utredningsförslag. Den här funktionen använder sig av Kungliga bibliotekets (KB:s) språkmodell KB-BERT som tränats på stora mängder svensk text. Syftet med applikationen är att underlätta för handläggare på statliga myndigheter att svara på nya utredningsremisser från regeringen genom att enkelt kunna hitta och ta inspiration från vad relevanta remissinstanser skrivit om tidigare, liknande förslag.

Applikationen är en prototyp som utvecklats av Ekonomistyrningsverket (ESV) i samverkan med Statens servicecenter och Tillväxtverket inom ramen för ESV:s datalabb.

Figur 1: Användargränssnitt för Hitta Remissvar

Hitta förslag och remissvar

Ska du svara på en remiss och vill ta stöd av vad din eller någon annan myndighet har svarat tidigare i liknande fall? Eller är du bara nyfiken på vilka förslag som lagts inom utredningsväsendet för ett visst område? Denna söktjänst hjälper dig att snabbt identifiera förslagen i statliga utredningar, hitta liknande förslag i andra utredningar och se vilka remissinstanser som lämnat synpunkter.

Observera att de allra nyaste utredningarna inte finns i söktjänsten eftersom prototypen ännu inte har automatisk uppdatering från källan.

Hjälp oss utveckla applikationen! I [denna enkät](#) kan du lämna synpunkter på befintlig tjänst och tycka till om framtida utveckling.

Sök på utredningsförslag

Utredning

<input type="text" value="T.ex. SOU 2022:15, Ds 2022:22 eller utredningens titel"/>	<input checked="" type="checkbox"/> SOU <input checked="" type="checkbox"/> Ds
<input type="text" value='T.ex. klimat eller "ekonomisk styrning"'/>	Sortera förslag Välj sortering <input type="text"/>
<input type="text" value="T.ex. Ekonomistyrningsverket eller Tillväxtverket"/>	Avgränsa år 2002 <input type="text"/> - 2023 <input type="text"/>

Träffar per sida: 20 0 av 0 < >

Demonstrationer

Ett antal live-demonstrationer (demos) av applikationen har genomförts. Dessa demonstrationer har inneburit att visa applikationens funktionalitet live för en publik. Under demonstrationerna har vi fått en hel del feedback som vi har noterat och i många fall tagit hänsyn till i vidareutvecklingen av applikationen.

Det finns flera syften och fördelar med att hålla demonstrationer av en applikation under utveckling. Genom att hålla en demo får man ofta

- spontana reaktioner från nya användare
- input på upplevelser av gränssnittet
- synpunkter och frågor som vi själva kanske inte tänkt på eller frågat oss
- insyn från grupper som inte är representerade i utvecklingsgruppen
- input på möjliga datakällor.

Vi utförde följande typer av demonstrationer:

- **Frukostseminarium - Lansering av Hitta remissvar**
Denna demonstration genomfördes 9 mars 2023 digitalt och på ESV kontor med deltagare från en rad olika myndigheter.
- **Demonstrationer för olika myndigheter**
Dessa typer av demonstrationer hölls för ett antal anställda på Skatteverket, Pensionsmyndigheten, Transportstyrelsen, Tillväxtverket och ESV.
- **Presentation för eSam**
Applikationen presenterades under ett eSam-möte med den övergripande rubriken “Text som data i staten”. Presentationen hölls 8 juni 2023.
- **Demonstration på AI Sweden**, inom ramen för AI Swedens Vinnovafinansierade projekt Interdisciplinary expert pool for NLU.
Denna demonstration hölls 12 september 2023.
- **Styrgruppsmöte myndighetssamverkan inom utbildningssektorn**
Under hösten 2023 frågade Tillväxtverkets enhet för Kompetensförsörjning och Digitalisering om vi kunde presentera möjligheterna för offentliga aktörer att använda språkteknologi genom att bl.a. demonstrera Hitta remissvar på styrgruppsmötet för myndighetssamverkan inom utbildningssektorn. Detta möte hölls 22 september 2023.
- **Ekonomistyrningsverkets ekonomichefs- och controllerchefsråd**
Detta möte ägde rum 7 december 2023.

Generell feedback från demonstrationerna

Deltagarna på de olika demonstrationerna har generellt varit mycket positiva, både till själva initiativet att utveckla Hitta remissvar, tjänsten i sig samt arbetssättet att presentera en prototyp som ger andra statstjänstemän insikt i vilka möjligheter språkteknologi och AI för med sig. Det har också kommit många konkreta utvecklings- och förbättringsförslag. De som deltagit i demos och workshoppar är generellt personer som varit intresserade av projektet vilket antagligen färgar reaktionerna

Ett återkommande användarönskemål var att få sökresultaten grupperade efter utredning i stället för förslag, så man inte får upp 394 träffar när det rör 12 remissvar på grund av att varje utredning innehåller många förslag. I den uppdaterade versionen av Hitta remissvar kommer remissvaren att vara sökbara, vilket möjliggör att vi kan matcha delar av ett remissvar mot ett specifikt förslag, men den funktionaliteten finns inte i prototypen som demonstrerades. Även för en sådan funktion skulle det kunna vara relevant att gruppera resultat.

Det fanns flera funderingar och frågor under lanseringen av Hitta remissvar. Någon finner det starkt begränsande att remissvar från innan 2016 inte finns att hitta i appen och undrar ifall det finns planer på att tillgängliggöra äldre material. Någon tycker att

detta riskerar att bli ett demokratiproblem. Eftersom inga remissvar från tidigare än 2016 finns på regeringen.se så finns det inte någon tillgänglig och enkel metod för att hämta denna data på ett samlat och automatiskt sätt. Om alla diarieförda remissvar fanns digitalt, strukturerat och maskinläsbart skulle det medföra stora möjligheter att tillgängliggöra även äldre material.

En annan användare undrar hur vi har gjort för att hantera myndigheter som byter namn. Svaret är att man måste söka på båda namnen då vi endast använt oss av de namn på remissinstanser som anges som filnamn på regeringen.se där vi hämtar alla remissvar. Detta innebär att remissinstansernas namn kan vara felstavade, inkonsekvent huruvida det är hela myndighetsnamn eller förkortningar som används, samt att remissinstanser kan heta saker som Tullverket_webb, eller Kriminalvården_kopia, Linköpings kommun_slutlig och liknande.

Under en demo fick vi tips om att testa att använda en annan, nyare KB-BERT-språkmodell³⁵.

Ett förslag på fler möjliga lösningar som vi fick var att inuti Hitta remissvar även erbjuda en tjänst som skriver utkast till remissvar.

Förutom konkreta tips och önskemål har vi också fått mycket positiv respons. Flera personer diskuterade och föreslog även andra användningsområden för språkmodeller och datadriven texthantering på myndigheter. De gav uttryck för att detta projekt är ett steg närmre en mer datadriven offentlig sektor.

Enkät

Under sommaren och hösten 2023 har det funnits en möjlighet för alla användare av Hitta remissvar att bidra till utvecklingsarbetet genom att lämna synpunkter i en webbenkät. 39 personer svarade på enkäten och det är framför allt myndighetsanställda med erfarenhet av att svara på remisser och liknande arbetsuppgifter. Det är troligt att de som svarat på enkäten är intresserade av den här typen av nya arbetsverktyg. Vi kan därför inte anta att de utgör ett representativt urval av möjliga användare.

Alla användare som svarat på enkäten såg nytta med Hitta remissvar för sina arbetsuppgifter. Möjligheten till att hitta underlag ses som den största nyttan. Tidsbesparingen per person uppskattas till enstaka timmar per månad och den är svår att uppskatta för de som svarat på enkäten. Slutligen ser användarna många utvecklingsmöjligheter för ett liknande verktyg. Några saker som användarna

³⁵ I vår applikation använder vi språkmodellen KB-BERT för att kunna matcha förslag mot varandra med hjälp av semantik. Språkmodellen är tränad på stora mängder svensk text. Se detaljerad beskrivning i Malmsten, M., Börjeson, L., och Haffenden, C. (2020), [Playing with Words at the National Library of Sweden—Making a Swedish BERT](#).

efterfrågade var att kunna gruppera resultat efter utredning i stället för resultat och att få tydligare kontext och översikt. Några användare förstår inte nuvarande funktionalitet fullt ut och flera har lyft nackdelar och risker med den begränsade datatillgången.

I följande avsnitt redogör vi detaljerat för resultaten i enkäten, som bestod av 10 frågor. För varje fråga presenterar vi resultaten i en tabell, förutom för de frågor där endast fritextsvar var möjligt. Resultaten beskrivs även i textform. För de frågor där fritextsvar har varit möjligt ger vi en sammanfattning av dessa samt viss diskussion.

Vi har valt att presentera könsuppdelad statistik för fråga 2 och 6. Observera att det endast är 12 män och 22 kvinnor som svarat på enkäten. Övriga deltagare har valt alternativet ” Vill ej ange” eller har lämnat frågan om könstillhörighet tom.

Fråga 1: I vilka arbetsuppgifter skulle du kunna använda tjänsten?

Tabell 1. Fråga 1 om arbetsuppgifter

Alternativ	n	Procent
I arbetet med att besvara remisser	33	82,5 %
I arbetet med att bereda beslutsunderlag	28	70,0 %
I arbetet med utvärderingar, analyser m.m.	25	62,5 %
Annat, nämligen:	4	10,0 %
Tjänsten är inte till nytta i någon arbetsuppgift	0	0,0 %
Vet ej	0	0,0 %

Användarna fick välja flera alternativ. Vi fick 3 fritextsvar.

Resultat:

- Alla som har svarat på enkäten verkar tycka att appen fyller någon form av nytta.
- Den vanligaste uppgiften som appen underlättar är enligt användarna att besvara remiss. 82,5 procent valde detta alternativ. Den näst vanligaste uppgiften som är att bereda beslutsunderlag var det 70 procent som valde.

Utöver de uppgifter vi själva har listat som exempel för vad appen skulle kunna underlätta nämner användarna även forskning, generellt utredningsarbete samt bevakning av särskilda ämnen såsom kommande förenklingar för företaget.

Fråga 2: Hur användbar är denna tjänst för de arbetsuppgifterna? 0 = inte alls och 10 = alltid användbar

Tabell 2. Fråga 2 om användbarhet

Alternativ	n	Min-värde	Max-värde	Medel-värde	Median	Standard-avvikelse
Kvinnor	20	0	10	6,1	6,5	2,9
Män	12	0	10	6,1	6,0	2,6
Vill ej ange	4	0	10	7,0	7,0	0,8
Totalt	36	0	10	6,3	7,0	2,6

Resultat:

- Medelvärde på 6,3 och medianvärdet på 7 visar på att appen är användbar för de nämnda uppgifterna, men inte maximalt användbar. Därför är fortsatt utveckling av applikationens funktion och gränssnitt viktig för att kunna öka användbarheten.
- Kvinnor och män skiljer sig inte nämnvärt åt vad gäller hur användbar de tycker att tjänsten är för de nämnda arbetsuppgifterna.

Fråga 3: Vilka fördelar/nyttor bidrar tjänsten med för dig?

Tabell 3. Fråga 3 om nyttor och fördelar

Alternativ	n	Procent
Jag sparar tid	28	71,8 %
Jag kan hitta underlag jag annars inte hade identifierat	36	92,3 %
Jag kan hitta underlag jag känner till, men inte visste var de fanns	24	61,5 %
Jag kan få perspektiv och infallsvinklar jag inte tänkt på innan	30	76,9 %
Resultatet av mitt arbete blir bättre	22	56,4 %
Annat, nämligen	3	7,7 %
Vet ej	0	0,0 %

Användarna fick välja flera alternativ. Vi fick 3 fritextsvar.

Resultat:

- Att hitta underlag man inte hade hittat utan applikationen är den största nyttan för användarna. Hela 92,3 procent har valt detta alternativ. Att få nya perspektiv och infallsvinklar samt att man sparar tid är också nyttor är också ett alternativ som många valt.
- Alla de nyttor vi föreslagit håller mer än 50 procent av användarna med om.
- Förutom de föreslagna nyttorna nämner användarna kvalitetssäkring och mer kunskap som nyttor och datadrivenhet som en fördel.

Fråga 4: Hur mycket tid kan du spara med hjälp av tjänsten?

Tabell 4. Fråga 4 om tidsbesparing

Alternativ	n	Procent
Upp till en timme per månad	4	14,8 %
1-2 timmar per månad	8	29,6 %
3-6 timmar per månad	5	18,5 %
Mer än 6 timmar per månad	1	3,7 %
Vet ej	9	33,4 %

Resultat: Av de svarande anger 33 procent att de har svårt att uppskatta hur mycket tid de kan spara med hjälp av applikationen, medan 30 procent tror sig kunna spara 1 till 2 timmar per månad. Att en stor del av användarna har valt alternativet ”vet ej” pekar på att frågan om tidsbesparing är svår att svara på.

Vår utformning av frågan och framför allt dess alternativ påverkar sannolikt de svar vi får. Mittenalternativen är de svar som flest användare valt.

Fråga 5: Ser du några nackdelar, problem eller risker i att använda denna tjänst

Fråga 5 var en fritextfråga där vi fick 20 fritextsvar, som vi summerar här.

Resultat:

- Tidsbegränsningen vad gäller inkludering av endast remissvar bak till 2016 beskrivs som ”farligt” av en användare och kritiseras av flera. Någon nämner dock att detta är bättre än inget.
- Flera användare hade svårt att förstå varför vissa matchningar var högre upp i resultatlistan än andra i de fall då användaren får resultat baserat på en semantisk sökfunktion snarare än nyckelord. Detta sker då användaren väljer att leta efter liknande förslag. Då man gör en nyckelord-sökning markeras nyckelordet i resultaten. Detta verkar vara något som användare gärna hade sett även för den semantiska sökfunktionen.

- En användare nämner risken att förlita sig på information man inte läst igenom helt, utan kanske bara de utdrag som appen presenterar. Som applikationen ser ut i nuläget är det dock möjligt att se hela kontexten med ett knapptryck. Även hela dokumentet finns tillgängligt i form av en pdf, även denna ett knapptryck bort.
- Någon nämner att det är lite övermäktigt att få många utdrag från samma utredning, utan hen hade hellre fått träffar på endast hela utredningar.
- En användare var missnöjd med att vissa SOU:er och Ds:ar inte fanns med, men skrev att detta kanske har att göra med att de är för nya.
- En person kände sig osäker på hur säker hen kunde vara på att de resultat hen får är korrekta.
- Ifall man får 0 sökträffar kan man luras att tro att det inte finns några relevanta dokument på ämnet, fastän resultatet kanske beror på att dessa dokument bara inte finns i tjänsten, eller att de hade dykt upp ifall man hade genomfört sin sökning på ett annorlunda sätt.
- En person tycker att appen är bra och skriver att det är en allmän rättighet att enkelt få tag på information.

Fråga 6: Vilken ytterligare funktionalitet skulle göra tjänsten mer användbar för dig?

Tabell 5. Fråga 6 om ytterligare funktionalitet

Alternativ	n, Total	Procent, Kvinnor, n= 22	Procent, Män, n=12	Procent, Vill ej ange, n= 4	Procent , Total
Möjlighet att få en sammanfattning av och/eller söka i konsekvensutredningar	21	45,5 %	66,7 %	75,0 %	53,8 %
Möjlighet att söka i och sortera ut just de delar av ett eller flera remissvar som handlar om ett specifikt förslag	26	72,7 %	50,0 %	100,0 %	66,7 %
Möjlighet att i en samlad översikt se om remissinstansen är huvudsakligen positiv eller negativ inställd till en specifik utredning eller förslag	16	45,5 %	25,0 %	75,0 %	41,0 %
Möjlighet att få en kort sammanfattning av remissvaren	23	63,6 %	58,3 %	50,0 %	59,0 %
Möjlighet att göra sökningar i andra texter, tex publicerade analyser och utredningar från myndigheter	24	59,1 %	58,3 %	100,0 %	61,5 %
Möjlighet till bättre användargränssnitt	9	22,7 %	16,7 %	50,0 %	23,1 %
Annan funktionalitet eller förbättring, nämligen	8	18,2 %	8,3 %	50,0 %	20,5 %

Användarna fick välja flera alternativ. Vi fick 8 fritextsvar, som vi summerar här.

Resultat:

- Att kunna söka i remissvaren är det mest populära alternativet.
- Att kunna söka bland andra dokumenttyper är näst mest populärt. Man vill gärna kunna söka bredare än bara SOU:er och Ds:ar.
- Fler kvinnor än män bedömer de ytterligare funktionerna som användbara i alla fall förutom möjligheten att få en sammanfattning av och kunna söka i konsekvensutredningar.
- Begränsad tillgång till äldre material nämns än en gång som ett problem.
- Någon undrar hur man exkluderar ett ord. Detta framgår av applikationens hjälpsida. Det kan alltså vara värt att tänka ett extra varv på hur vi presenterar de instruktioner vi har för hur man använder appen. Ett alternativ är att ha en knapp precis bredvid sökfältet som tar användaren direkt till hjälpsidan, alternativt presentera ett utdrag ur instruktionerna direkt under sökrutan på startsidan, samt att utveckla tydliga FAQ + instruktionsfilmer.
- Någon vill gärna ha en chattfunktion där chatteboten baserar sina svar på olika typer av myndighets- riksdags- och EU-dokument.
- En annan person vill gärna ha en mobilanpassad version av appen.

Fråga 7: Övriga kommentarer om tjänsten

Användarna hade också möjlighet att lämna övriga kommentarer om tjänsten i enkäten. Vi fick 14 fritextsvar som vi summerar här.

Resultat:

- Ett problem med att svara på enkäten beskriver en användare som att hen inte har använt appen i ”skarpt” läge, utan endast i undersökande syfte. Detta påverkar antagligen personens svar. Detta kan vara något som flera enkätbesvarare upplever. Att lägga till en fråga om hur många gånger användaren testat appen skulle kunna hjälpa oss att förhålla oss till detta.
- En person nämner att den gärna vill se att förslag från samma utredning sorteras ihop.
- En annan hade gärna sett ett API att kunna använda appen genom.
- Att kunna få en helhetsbild av vad som publicerats inom ett visst område beskrivs av en användare som viktigare än att kunna se och lista remissvar till särskilda utredningar.
- Någon tycker att presentationen av resultat är svår att ta sig an, att det är rörigt och för mycket text. Detta är något vi tar med oss i vidareutvecklingen av framför allt användargränssnittet.

En person skulle gärna se en funktion där man får veta hur stor procent av remissinstanserna som inte har lämnat några synpunkter alls.

Fråga 8: Ungefär hur många remissvar arbetar du med per år?

Tabell 6: Fråga 8 om antal remissvar per år.

Alternativ	n	Procent
0	2	5,1%
1-2 remissvar	11	28,2%
3-4 remissvar	9	23,1%
5 eller fler remissvar	17	43,6%

Resultat: 40 procent av de som svarat på enkäten svarar på 5 eller flera remissvar om året. Bara 2 personer verkar inte svara på remisser alls.

Fråga 9: Hur många år har du arbetat med utredningar och remisser?

Tabell 7. Fråga 9 om erfarenhet av utredningar och remisser

Alternativ	n	Procent
Mindre än 1 år	4	10,2%
1-5 år	12	30,8%
Mer än 5 år	23	59,0%

Resultat: Nästan 60 procent av de som svarat på enkäten har arbetat med utredningar och remisser i mer än fem år.

Fråga 10: Är du kvinna eller man?

Tabell 8: Fråga 10 om könstillhörighet³⁶

Alternativ	n	Procent
Kvinna	22	57,9%
Man	12	31,6%
Vill ej ange	4	10,5%

Resultat: Fler kvinnor än män har svarat på enkäten.

³⁶ I anslutning till frågan fanns följande förklaringstext: "I Sverige finns i dagsläget två juridiska kön, kvinna och man. Det utesluter inte att personer identifierar sig på annat sätt. Anledningen att vi gör denna insamling är för att kunna rapportera statistik för hur insatser når kvinnor respektive män, men även som underlag för att analysera om det finns stora skillnader hur tjänsten uppfattas."

Förenkling av användargränssnitt

En UX-designer fick i uppgift att ge förslag på hur gränssnittet skulle kunna göras mer användarvänligt. Han föreslog följande förändringar:

En enkel sökruta. Till höger om resultatsidan presenteras filtreringsmöjligheter.

Detta förslag är menat att förenkla för användaren. Denna version bör vara mer intuitiv då de stora sökmotorerna oftast bara har en sökruta. Att anpassa gränssnittet efter användarens vanor innebär alltså en förenkling. Att lägga filtreringsalternativ ett steg senare, alltså när användaren har gjort en sökning, visar på den hierarki som finns för de rutorna som användaren kan fylla i då filtreringen är underordnad sökfunktionen.

När du öppnar applikationen och inte har gjort en sökning än visas instruktioner för hur du kan anpassa din sökning med hjälp av variabler. Dessa instruktioner är ett utdrag från den information som finns att läsa under ”hjälp”-fliken i appen. Genom att presentera instruktionerna i start-vyn underlättar vi för användaren att kunna anpassa sin sökning utan att behöva leta efter hjälp själv.

Texten som beskriver appen bör vara så kort och koncis som möjligt

Om denna information är för lång kommer användaren antagligen bara att hoppa över att läsa denna.

Rullgardinsmeny för remissinstanser

I stället för att själv skriva in remissinstans som första val föreslås en drop-down-menü för möjliga remissinstanser. Ifall man väljer att endast ha en sökruta för remissinstans så bör exempel på typer av instans ges i sökrutan (myndigheter, företag, föreningar, etc.).

Detta förslag ämnar förenkla processen på så sätt att användaren inte behöver riskera att stava fel, och att myndighetsnamn är normaliserade. Det senare skulle kräva viss databehandling då vi i nuläget indexerar namnen för remissinstanserna baserat på de filnamn som regeringen.se använt. Det ska fortfarande vara möjligt att kunna skriva in en remissinstans själv alternativt söka efter den instans man är intresserad av i drop-down-menyn. Problemet med detta förslag är att instanserna inte är begränsade för myndigheter, utan inkluderar olika företag och föreningar som har svarat på remiss. Detta gör att vi inte kan använda oss av en generisk lista av alternativ för drop-down-menyn, utan denna skulle behöva uppdateras baserat på vilka utredningar som visas under ”sökresultat”. Denna komplexitet är dock något som UX-designern inte hade uppfattat/fått information om. Ifall man väljer att endast ha en sökruta för remissinstans på grund av denna komplexitet skulle man ändå kunna överväga att

göra en insats för att automatiskt normalisera myndighetsnamn, alltså att se till att det inte finns flera namnvariationer för en och samma myndighet.

Termen "sökresultat" borde bytas ut mot "förslag"

Detta skulle förtydliga vad det faktiskt är användaren får fram, samt vad antalet sökträffar egentligen syftar på. Till exempel skulle användaren kunna bli osäker på ifall det är antal utredningar eller förslag som åsyftas.

Ikonerna som finns för att "fälla ut" flikar i applikationen bör vara enhetliga

När användare ska fälla ut förslag (då man innan bara ser en rubrik) visas ikonen "+". Då visas början av förslagets text. Man kan sedan använda "-" för att gå tillbaka till att endast se förslagets rubrik. När man har fällt ut förslaget första gången med "+"-ikonen får man sedan alternativet att fälla ut förslaget ytterligare, för att kunna se förslagets fullständiga text. Detta visas dock med en annan ikon, en pil som pekar nedåt. Man kan sedan fälla in förslaget igen genom att klicka på nederdelen av texten. UX-designerns förslag är att använda sig av samma typ av ikon för alla gånger då användaren kan fälla ut förslag. Dessa ikoner bör vara pilar.

Det borde gå att göra en sökning genom att endast klicka på "Enter" på tangentbordet

I appen söker man genom att klicka på ikonen "sök". Ifall användaren klickar på "Enter" på tangentbordet händer ingenting. Eftersom det är standard att kunna söka med "Enter" för sökmotorer vore även denna funktion intuitiv då det är en etablerad vana.

Möjlighet att klicka sig vidare

Användaren bör kunna "klicka sig vidare" efter en sökning snarare än att få en extra spalt för liknande förslag där sökmöjligheterna tar slut. Ett exempel följer: Användaren söker på SOU X och denna dyker upp i sökresultaten, med dess förslag listat under. Användaren klickar på ett förslag från SOU X. Då presenteras detta förslag från SOU X i högerspalten, under filtreringsmöjligheterna. Om användaren då väljer klickar på "visa liknande förslag" visas liknande resultat i vänsterspalten. Användaren ska kunna backa i sin sökhistorik genom att backa i något som ser ut som ett filtrad som finns ovanför sökrutan. Det bör också finnas möjligheter att backa i historiken med en bakåt-pil. Detta förslag ämnar göra det lättare att kunna ångra sig, att alltså kunna backa när som helst. Det öppnar också för att kunna fortsätta att välja liknande förslag, och alltså ha möjlighet till att klicka sig vidare "oändligt".

Tydligare information om sortering

Det borde synas vilken sortering som gäller för ens sökresultat. Ifall man gör en nollsökning är det oklart vilken typ av sortering som gäller: fallande, stigande eller

relevans. Det borde synas i gränssnittet vilken typ av sortering som är default, alltså standardinställningen.

Workshoppar med utredare för att testa sammanfattningar av remissvar och andra funktioner

Workshop 1

I denna workshop deltog 4 utredare samt 3 representanter från arbetsgruppen. Innan workshoppen hade de fått information om att de skulle utvärdera ny möjlig funktionalitet för Hitta remissvar.

Workshoppen utgick ifrån en särskild utredning, och tolv olika remissvar till denna. Vi valde också ut ett specifikt förslag från utredningen. Vi tittade på SOU 2020:34 Stärkt kvalitet och likvärdighet i fritidshem och pedagogisk omsorg. ”

Deltagarna fick vid starten av workshoppen veta vilken utredning, vilket förslag och vilka remissvar som skulle undersökas. För denna funktionalitet finns inte något användargränssnitt utvecklat, utan vi testade direkt i en s.k. programutvecklingsmiljö.

Vi hann under workshoppen inte gå igenom samtliga tolv remissvar.

Syftet med workshoppen var att undersöka hur väl de generativa språkmodellerna GPT-4 och GPT-SW3 utförde följande uppgifter gällande varje remissvar:

1. Bedömning av ifall remissvar har synpunkter på hela remissen.
2. Bedömning av ifall remissvar har synpunkter på det specifika förslaget (och i så fall var i remissvaret).
3. Bedömning av ifall de generella synpunkterna är positiva, negativa eller neutrala.
4. Bedömning av ifall synpunkterna på det specifika förslaget är positiva, negativa eller neutrala.
5. Sammanfattning av synpunkter på det specifika förslaget.

Huvudsakliga resultat:

Båda modellerna lyckades i samtliga testade fall fånga om det fanns ställningstaganden till utredningen i remissvaret. GPT-SW3 ansåg nästintill alltid att remissinsatserna var positivt inställda till hela utredningen, där var GPT-4 något mer balanserad. Kopplat till detta, hade vi en diskussion om det ens är möjligt att klassificera hur remissvar ställer sig till en hel utredning då de ofta är mer komplext än alternativen positiv, negativ eller neutral. Medskicket var att vid nästa tillfälle testa med fler kategorier såsom ”överlag positiv/överlag negativ” samt ”osäker”.

Modellerna var i vissa fall bättre än människan (som i vissa fall bara hade hunnit läsa in sig på sittande möte) på att identifiera om det fanns synpunkter i remissvaret kopplat till ett specifikt förslag. Där hittade båda modellerna ytterligare synpunkter på sidor i remissvaret som människan inte identifierade.

GPT-SW3 var tydligt sämre än GPT-4 på att sammanfatta remissinsatsens åsikt om förslaget. I två av tre körningar kom inte ens remissinstansens ställningstagande till förslaget fram i sammanfattningen.

I slutet hade vi en diskussion om hur man ska hantera genererade sammanfattningar då ordval och kontext kan vara essentiella för remissvarets åsikter och dessa riskerar att falla bort i genererade sammanfattningar. Här fördes en diskussion om att eventuellt också ha en funktion där hela remissvarets text om ett specifikt förslag kan återges i text.

Något som vi upptäckte under workshop 1 var att det inte behöver vara så många utredare som deltar. Det är viktigare att testa vid fler tillfällen än att vara fler som testar vid ett tillfälle. Om man är ett färre antal personer hinner man med mer och det blir enklare att genomföra uppgifterna.

En annan allmän synpunkt från flera utredare var dock att det var värdefullt för dem att vara med och testa i det här skedet och den här formen, eftersom det gav förståelse för hur språktekniken och AI fungerar i praktiken.

Workshop 2

I denna workshop deltog 2 utredare samt 4 representanter från arbetsgruppen. Vi tog till oss av lärdomar och rekommendationer från den första workshoppen och valde därför att ha en mindre grupp denna gång.

Ungefär samma upplägg som användes i workshop 1 användes även för workshop 2, med vissa förändringar. Utredarna fick chansen att förbereda sig då ett utskick gjordes med info om specifikt förslag, remissvar samt de frågor vi planerade besvara under workshoppen. Vi gjorde ingen sentimentanalys på hela förslaget. Under workshoppen hann vi titta på två olika förslag från två olika utredningar.

Vi valde att titta på ”svårare” remissvar – därav fick vi kanske ett mer negativt resultat än det hade kunnat vara. Vi hann inte gå igenom samtliga remissvar under workshoppen. De utredningar vi utgick ifrån var SOU 2020:48 Skatt på engångsartiklar, och SOU 2021:95 En möjlighet till småskalig gårdsförsäljning av alkoholdrycker.

Huvudsakliga resultat:

Modellerna (eller i alla fall GPT-4) bedömer ofta att flera olika sidor innehåller synpunkter om det specifika förslaget. Detta har till viss del att göra med att vi behövde dela upp remissvaren i sidor och fråga GPT ifall sidan innehöll synpunkter på förslaget eller ej. Eftersom kontextlängden för modellerna är begränsad var detta nödvändigt. Detta medförde alltså att vi inte kunde be GPT att endast välja en av sidorna som ”mest trolig att innehålla synpunkter” då varje sida behandlades som ett separat dokument för en separat uppgift. De sidor som GPT-4 markerat att innehålla synpunkter täcker ofta in den sida som annoterats av en människa som innehållande av synpunkt, men GPT-4 pekar också på andra ställen i remissvaret. När dessa ”falska positiva” sidor har gått igenom har det inte varit helt enkelt för utredare att bedöma ifall dessa sidor faktiskt innehåller åsikter om det specifika förslaget eller ej. Det verkar alltså vara svårt för både människa och maskin att göra en bedömning i många fall.

Sentimentanalys var heller en inte helt enkel uppgift då remissinstanser ofta är positiva till den övergripande inriktningen och syftet, men negativa till den specifika utformningen av förslaget. Ett exempel: Man är positiv till hållbarhet och att utredningar görs kring detta tema, men tycker att ett eller flera förslag är dåliga eller problematiska på något sätt. Det är svårt att göra en sentimentanalys även manuellt.

GPT-4 är generellt bra på att sammanfatta synpunkter på förslag (detta efter att vi själva manuellt valt de sidor som ska sammanfattas). GPT-SW3 är alla gånger sämre än GPT-4 på denna uppgift. Ibland gör GPT-SW3 ett okej jobb, andra gånger ett dåligt jobb. En bidragande faktor kan vara att vi kanske är bättre på att prompta GPT-4 än GPT-SW3.

Sammanfattningarna som skapades av GPT-4 gav oftare en bredare bild än synpunkterna på själva förslaget. Detta var dock något som generellt bedömdes som positivt. GPT-4 började oftast sammanfattningarna med det som var mest relevant för förslaget och gick sedan vidare till en bredare bild som gav kontext till förslaget. Detta är spännande med tanke på att synpunkten på det specifika förslaget inte alltid presenterades först i de sidor som GPT-4 ombads att sammanfatta.

Vi testade att sammanfatta samma remissvar med GPT-modellerna flera gånger för att kontrollera hur pass konsekventa de var. GPT-SW3 ger ofta väldigt olika sammanfattningar trots samma prompt, även i de fall då inga parametrar hade ändrats. En gång skrev den t.ex. att instansen tyckte att den föreslagna skatten var bra, en annan gång skrev det inte var lämpligt att införa skatten. Detta kan till viss del ha att göra med att vi använde oss av en högre temperatur (ska avgöra hur pass "kreativ" modellen är) för GPT-SW3 då vi bedömde att GPT-SW3 behövde en högre temperatur för att generera sammanhängande text.

En möjlig väg framåt är att man skulle kunna sortera sidorna beroende på hur väl de matchar det specifika förslaget istället för att som i nuläget använda sig av binär klassificering med hjälp av GPT (klassificeringen av om en sida innehåller en synpunkt på detta förslag: ja eller nej).

Den nya lösningen kan göras med en semantisk vektorsök-modell, vilket redan används i andra delar av nuvarande Hitta remissvar-applikationen. På så sätt får man en "sannolikhet" för varje sida i stället för en snäv binär klassificerare där man inte får någon sortering av de olika sidorna vad gäller hur sannolikt det är att de beskriver det särskilda förslaget.

Vi drog slutsatsen att det stämde att det var tillräckligt med några få utredare som deltog i workshopen.

Skillnader workshop 1 och 2

Under första workshopen tyckte vi att både GPT-4 och GPT-SW3 var bra på att bedöma ifall remissinstansen hade synpunkter eller ej på utredningsnivå. Under andra workshopen med tuffare exempel tyckte vi att GPT-4 var överlägsen och GPT-SW3 ofta var inkorrekt.

Workshop 3

Under workshop 3 testade vi funktioner som är planerades inför den då kommande versionen av Hitta remissvar. I denna workshop deltog 4 personer från arbetsgruppen och 6 utredare.

Under denna workshop en ny funktion som innebär att kunna matcha ett förslag i en remiss mot utdrag från remissvar. Användaren kan antingen välja att söka bland alla remissens remissvar, eller söka i utvalda remissvar och hitta det mest relevanta stycket ur detta.

De utredningar som undersöktes under workshopen var SOU 2021:68, SOU 2022:39, SOU 2023:20, SOU 2021:81 och SOU 2022:49.

Huvudsakliga resultat:

Det fördes en diskussion om relevans och vad det innebär i denna kontext. För att tydligare beskriva vad funktionen faktiskt gör är att i stället för att kalla det "relevans" istället prata om "likhet", vilket är vad funktionen baserar sina matchningar på. Denna terminologi kan vara viktig främst för beskrivningen av applikationen och dess hjälp-sida.

Ett sätt att tydliggöra vad resultaten faktiskt baseras på är att referera till ”skrivelser” eller ”text” snarare än ”synpunkter”. All text i remissvaren är ju nödvändigtvis inte synpunkter, utan kan vara referenser, kommentarer, förtydliganden, etc.

Det är eventuellt problematiskt att stycken ur remissvar som citerar remissens förslag hamnar högst eller högt upp i resultatlistan. Det fördes en diskussion om ifall man borde ta hänsyn till detta på något sätt: t.ex. exkludera stycken som till större delen endast innehåller citat från remissen. Detta skulle dock kunna vara problematiskt då citaten är en stark indikation på att kringliggande text kommenterar det förslag användaren är intresserad av. Stycken med intressant text skulle därför försvinna. I många fall verkar rankningen av resultat fungera bra, men i vissa fall bedöms rankningen som icke-maximal av användarna.

Vi diskuterade även kring varför vissa remissinstansers svar hade hamnat högre upp än andra. En möjlig förklaring som togs upp var att vissa myndigheter är mer benägna än andra av att vara tydliga med exakt vilket förslag de faktiskt svarar på vart i sitt remissvar.

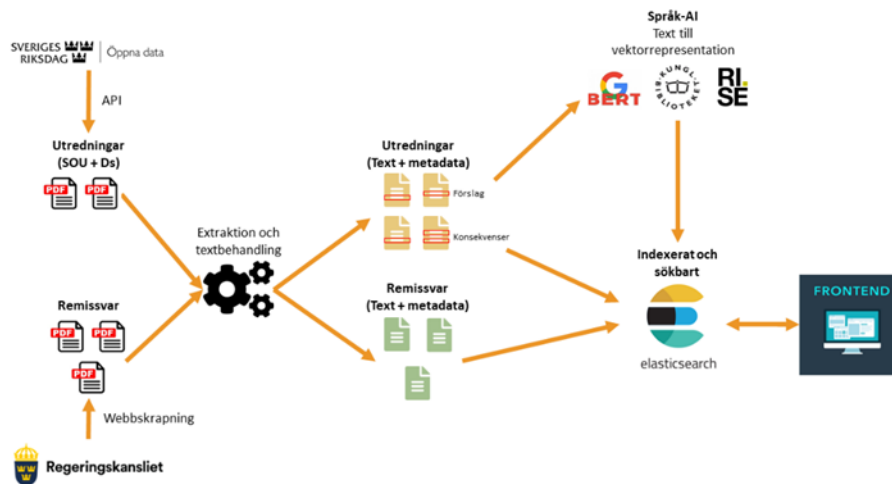
Vid några tillfällen hamnade stycken som var mindre relevanta från samma remissvar som ett stycke som rankas längre ner i resultatet. I ett fall hade regeringen publicerat själva remissen som ett remissvar till sig själv. För förslag ur denna remiss blev rankningen av stycken ur remissvar inte särskilt bra då stycken ur själva remissen fyllde ut resultatlistan.

Användare påpekade att för vissa remisser och dess förslag är kanske vissa specifika myndigheters svar mest intressanta, oberoende vad de faktiskt skrivit. En annan användare ville gärna få en "sammanfattning" av remissvar. Om man har fått ett remissvar som är högt rankat så kan man hitta till det mest relevanta stycket själv.

En användare uttryckte en önskan att kunna söka efter nyckelord i alla remissvar för att kunna hitta relevant underlag när hen ska svara på remiss. Någon nämnde att det skulle vara trevligt om applikationen hade mer grafiska inslag. En person nämnde att det är viktigt med appar likt Hitta remissvar för att få bättre institutionellt minne inom staten

Bilaga 3 – Teknisk beskrivning av prototypen

Figur 1: Hitta remissvars arkitektur



Hitta remissvar körs helt lokalt på maskiner i Ekonomistyrningsverkets produktionsmiljö. Indexeringen sker på en maskin med en kraftig grafikprocessor (GPU) men skulle kunna göras långsammare utan större besvär på billigare hårdvara.

Koden för projektet Hitta remissvar (internt namn: RemissAI) består av fyra delar:

RemissAIGUI

RemissAIGUI är koden till webbgränssnittet. Det är skrivet i Angular 17, ett Typescript-ramverk för webbutveckling. Även Bootstrap och jQuery används.

Eftersom webbgränssnittet bygger på Angular används flera komponenter för att visa informationen. För att minimera arbetet samt maximera möjligheterna har många komponenter och vyer återanvänts i olika kontexter. Alla sökparametrar som kan specificeras i gränssnittet speglas i webbadressen för att göra det möjligt att spara och dela sökningar.

Webbgränssnittet kommunicerar med servern genom ett API.

RemissAIWebAPI

RemissAIWebAPI är koden till API:t. Det är ett REST API skrivet i .NET 8.0 och exponerar data för appens alla behov. Datat lagras med hjälp av Elasticsearch. I Elasticsearch-instansen sparas dokumentens text, metadata, men även vektorrepresentationer för texten uppdelad i bitar, vilket möjliggör blixtnabba matchningar mellan olika objekt.

RemissAIWebscraper

RemissAIWebscraper är koden till en så-kallad scraper som är byggd i Python med hjälp av ramverket Scrapy. Denna scraper är kapabel att hämta hela webbsidor från regeringen.se och att spara nyckeluppgifter från dem. Den skapar en lista av statliga utredningar och dess remissvar med direktadresser för att kunna ladda ner dem.

RemissAIDataPipelines

RemissAIDataPipelines är koden till en annan del av projektet som gör den tunga data-bearbetningsprocessen i Python. Den gör flera saker:

- hämtar statliga utredningar som PDF från Riksdagens öppna data för att komplettera uppgifterna vi får från Regeringskansliets webbplats
- hämtar remissvar som PDF från regeringens webbsida genom katalogen som skapades av RemissAIWebscraper
- extraherar texten från PDF:erna med hjälp av specialiserade ramverk som PyMuPDF och OCRmyPDF
- extraherar metadata från texten och klassificera de olika delarna (rubriker, förslag, konsekvenser...) med hjälp av regelbaserade filter och ramverk som OpenCV
- anonymiserar texten genom att identifiera och ta bort e-postadresser, telefonnummer och även personnamn med hjälp av Kungliga Bibliotekets modell för namnigenkänning
- skapar inbäddningar för de olika delarna av utredningen och remissvarens text, dvs. en vektorrepresentation av dessa texters innehåll
- lagrar allt detta i Elasticsearch

Alla dessa steg körs automatiskt och regelbundet för att se till att Hitta remissvar visar aktuella data.

Övriga delar

Utöver dessa fyra delar har datalabbet skrivit mer kod för att utforska olika sidospår. Ett var vidareträna en modell för att klassificera utredningen och remissvarens innehåll. Ett annat var att använda stora språkmodeller såsom GPT-4 och GPT-SW3 genom OpenAI och AI Sweden:s API för att klassificera remissvarens innehåll.³⁷ För att göra detta har vi använt OpenAI:s API och Tillväxtverkets tillgång till AI Sweden:s beräkningsresurser, genom medverkan i ett Vinnova-finansierat projektet för att testa modellens kapaciteter.

³⁷ Vinnova, *GPT-SW3: en svensk basmodell för texthantering*, <https://www.vinnova.se/p/gpt-sw3-en-svensk-basmodell-for-texthantering/>. Uppdaterad 30 september 2022. Hämtad 26 februari 2024.

ESV gör Sverige rikare

- Vi har kontroll på statens finanser, utvecklar ekonomistyrningen och granskar Sveriges EU-medel.
- Vi arbetar i nära samverkan med Regeringskansliet och myndigheterna.