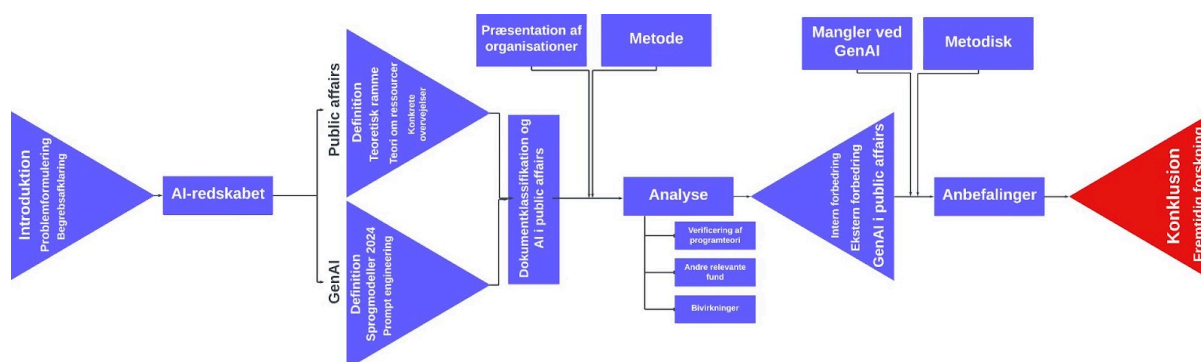


11.0 Konklusion og fremtidig forskning

Dette speciale har undersøgt problemformuleringen:

Hvilket potentiale er der i at anvende generativ AI – specifikt brugen af en sprogmodel til dokumentklassifikation – til at forbedre organisationers arbejde med public affairs?

Konklusionen er struktureret i 3 dele: Først en opsamling og konklusion på programteorien, dernæst en konklusion som svar på problemformuleringen ophængt i empirien fra evalueringen, indsigterne fra den konceptuelle kontekst og pointerne fra implikationsafsnittet og slutteligt er der bud på hvad fremtidig forskning kunne undersøge yderligere vedrørende GenAI i public affairs.



11.1 Opsamling på programteorien

Evalueringen af AI-redskabet i fire organisationer, SJ&K, Zero Carbon Shipping, Dansk Standard og Danske Gymnasier, har givet indsigt i, hvordan GenAI kan anvendes til at styrke public affairs arbejde, og hvor der stadig er behov for forbedringer i brugen af GenAI for at realisere potentialet. Helt konkret kan programteorien **ikke** bekræftes. Overordnet er grunden til nulfundet, at AI-redskabets monitoreringsrapport indeholder for mange falske positive, som medfører for meget støj, hvilket gør, at organisationerne ikke læser monitoreringsrapporten fast, hvorfor AI-redskabet dermed ikke kan medføre mere viden og bedre politiske og strategiske valg (DS1, SJ&K1, ZCS2).

Det kan også konkluderes at AI-redskabet kan være nyttig i små, specifikke kontekster, hvor falske positive ikke udgør et væsentligt problem. For Danske Gymnasier har AI-redskabet

ikke gjort en skelsættende forskel for deres arbejde med public affairs, men har stadig været nyttigt til at give dem indsigt i flere detaljer om hvad der foregår i Folketinget baseret på dokumenterne fra monitoreringsrapporten, og altså øget deres vidensniveau (DG3).

Metodeafsnittet fremhævede brugen af procesbaseret virkningsevaluering, som var afgørende for at forstå, hvordan og hvorfor AI-redskabet potentielt kunne skabe værdi. Denne tilgang tillod en undersøgelse af mekanismerne bag AI-redskabets funktion og identificerede specifikke områder for forbedring (Dahler-Larsen, 2018). Evalueringen af AI-redskabet på tværs af de fire organisationer afslørede udfordringer med falske positive og behovet for bedre prompt engineering. Disse udfordringer understøtter behovet for en iterativ udviklingsproces, hvor feedback fra brugerne aktivt inddrages for at optimere AI-redskabets præstationer eller andre fremtidige GenAI drevne værktøjer.

Nedenfor opsummeres hvad de 4 organisationer svarede og bidrog med af central empiri.

11.1.1 SJ&K

SJ&K benyttede AI-redskabet til at overvåge politiske dokumenter inden for 3 forskellige områder: tech, socialpolitik og fødevarepolitik, som en del af deres konsulent-virke. Evalueringen viste, at AI-redskabet havde potentiale til at skabe værdi ved at frembringe politiske indsigter, der kunne anvendes strategisk i deres public affairs arbejde, hvor især promptet for techvirksomheden gav gode resultater. Dog var der for mange falske positive, hvorfor SJ&K ikke læste monitoreringsrapporten fast (SJ&K1). Der var behov for en bedre sortering af indholdet i monitoreringsrapporterne, fordi organisationen ikke havde kapaciteterne til at bearbejde den store mængde information og heller ikke prioriterede det (SJ&K1). Omvendt var der to positive aspekter ved brugen af AI-redskabet i SJ&K: For det første er AI-redskabet et “sociologiske professionaliseringsværktøj” (SJ&K1), som gav en større bevidsthed om bevægelser i Folketinget, og blot det at modtage en monitoreringsrapport “er med til at man gør sig mere umage, og man taler om politiske beslutningsprocesser i Folketinget på en lidt mere kvalificeret måde.”, selvom monitoreringsrapporten ikke blev læst fast (SJ&K1). For det andet blev AI-redskabet brugt som en proxy for en RAG-model, hvor SJ&K bagudrettet læste igennem monitoreringsrapporter for at få et overblik over relevante politiske dokumenter i forbindelse med en public affairs arbejdsopgave for socialorganisationen (SJ&K1,5). Samlet set

vurderede SJ&K, at AI-redskabet bidrog til deres arbejde i en lille skala, men det blev også klart, at der var behov for bedre prompt engineering for at reducere mængden af falske positive og forbedre relevansen af de klassificerede dokumenter.

11.1.2 Zero Carbon Shipping

Zero Carbon Shipping, som er interesseret i maritim dekarbonisering, var primært inddraget i specialet som eksterne public affairs ekspert, men der blev stadig lavet en fuld evaluering af deres brug af AI-redskabet. På trods af deres begrænsede fokus på dansk politik fandt Zero Carbon Shipping, at AI-redskabet kunne levere nyttige indsigter i internationale politiske bevægelser og tendenser gennem dokumenterne fra Folketinget. Evalueringen afslørede dog udfordringer med implementeringen, især i forhold til promptsystemets præcision, mængden af irrelevante data og brugerfladen (ZCS1). Dog så Zero Carbon Shipping et potentiale ved AI-redskabet: “Og hvis jeg nu var en dansk aktør, så ville det være en helt anden sag. Så ville det jo være meget meget mere relevant.” (ZCS2). Qua Zero Carbon Shippings rolle som eksterne public affairs ekspert blev de som den eneste vist en demo af en RAG-model, som de omtalte som “overraskende godt” (ZCS2).

11.1.3 Dansk Standard

Dansk Standard var med som en public affairs organisation, der var en ‘least likely case’ givet det svære ved præcist at definere hvad der kan klassificeres som relevant for en public affairs organisation, hvor de fleste områder potentielt kunne være relevant i arbejdet med standarder (DS1). Dansk Standard brugte AI-redskabet til at overvåge standarder og politiske udviklinger inden for byggeri- og energisektoren. Evalueringen viste, at AI-redskabet kunne matche niveauet af dyrere og mere komplekse løsninger, hvilket selvsagt ikke er ensbetydende med en succesfuld implementering, men viser, at der er potentiale ved at anvende GenAI i public affairs til dokumentklassifikation (DS1). Udfordringen lå i håndteringen af den store mængde data og sikringen af, at informationerne var relevante og brugbare, hvor der var for mange falske positive (DS1). Dette understregede behovet for mere avanceret prompt engineering af AI-redskabet for at øge dets kvalitet og anvendelighed i en kompleks public affairs kontekst.

11.1.4 Danske Gymnasier

Danske Gymnasier var med som en public affairs organisation, der var en ‘most likely case’ givet at ungdomsuddannelser som genstandsfelt er tydeligt afgrænset og småt samt med min tilknytning og viden om Danske Gymnasiers arbejde givet min tidligere ansættelse der. Danske Gymnasier var den organisation, der gav den mest positive evaluering af AI-redskabet. Evalueringen viste, at AI-redskabet effektivt kunne identificere relevante dokumenter inden for et klart defineret genstandsfelt, hvilket lettede deres arbejde med at holde sig opdateret. Dog blev det klart, at AI-redskabet stadig producerede mange falske positive og også få falske negative, hvilket gik ud over tilliden til AI-redskabet. På trods af udfordringerne mente Danske Gymnasier, at AI-redskabet kunne forbedre deres evne til at være opdateret på politiske bevægelser i Folketinget, men det var ikke skelsættende for deres arbejde og derfor ikke direkte kunne hjælpe dem med at træffe bedre politiske og strategiske valg.

11.2 Konklusion på baggrund af konceptuelle kontekst

Problemformuleringen for dette speciale fokuserede på at undersøge potentialet ved at anvende generativ AI, specifikt brugen af sprogmodeller til dokumentklassifikation, for at forbedre organisationers arbejde med public affairs. Evalueringen af de fire organisationer viste, at AI-redskabet havde potentiale til at forbedre public affairs arbejdet på flere måder, men også at der var væsentlige implementeringsudfordringer, der skulle adresseres. Forbedringerne som GenAI kan bidrage med til public affairs kan sætte op mod den konceptuelle kontekst ved pluralismen og politisk monitorering som en katalytisk kapacitet som gennemgås nedenfor (Esbensen, 2007; Hegelund & Mose, 2012; Oberman, 2008).

Public affairs traditionen pluralisme, som fremhæver konkurrence mellem organisationer om indflydelse, understreger behovet for at være velinformeret om politiske bevægelser (Hegelund & Mose, 2012). AI-redskabet viste et potentiale ved at benytte GenAI i public affairs, som kunne være nyttigt i pluralismen ved at kunne analysere og klassificere store mængder politiske dokumenter, hvilket øgede organisationernes mulighed for at opnå en konkurrencefordel via øget viden om hvad der sker i Folketinget. Dette potentiale blev dog ikke fuldt ud realiseret, fordi AI-redskabet leverede for mange falske positive.

Politisk monitorering refererer til public affairs organisationers behov for konstant at være opdateret om politiske bevægelser og beslutninger (ibid). Denne evne til at mappe og scanne det politiske landskab via et GenAI værktøj kan ses som en katalytisk ressource (Oberman, 2008; Esbensen, 2007). Et GenAI værktøj som AI-redskabet kan forbedre denne evne på to måder: For det første kan GenAI værktøjer give organisationer mulighed for løbende at overvåge politiske dokumenter og hurtigt identificere de mest relevante informationer. Dette betyder, at organisationerne kan reagere hurtigere på politiske ændringer og justere deres strategier i overensstemmelse hermed. Danske Gymnasier fandt for eksempel, at AI-redskabet effektivt kunne identificere relevante dokumenter inden for deres klart definerede genstandsfelt, hvilket forbedrede deres evne til at holde sig opdateret (DG3). For det andet kan et GenAI værktøj reducere mængden af informationsstøj ved at benytte et finmasket prompt til at minimere mængden af irrelevante data og fokusere på de mest relevante politiske dokumenter. Dette gør det lettere for organisationerne at skelne mellem vigtig og uvæsentlig information, hvilket er afgørende for effektiv politisk monitorering (Hedelund & Mose, 2013). Alle fire organisationer oplevede dog at der var for meget informationsstøj stadig i monitoreringsrapporterne fra AI-redskabet, hvorfor potentialet for at forbedre public affairs skal findes ved en anden brug af GenAI end den måde GenAI er benyttet i AI-redskabet.

Sammenfattende konkluderer dette speciale, at AI-redskabet ikke lever op til programteorien, men via evalueringen af AI-redskabet blev der identificeret potentielle måder at forbedre public affairs arbejdet på ved at benytte GenAI til dokumentklassifikation, som kan gøre public affairs organisationer bedre til politisk monitorering. Dette potentiale kan udforskes yderligere i fremtidig forskning. Fremtidig forskning skal tage højde for det potentielle mismatch der kan ske mellem ressourcestærke public affairs organisationer, der kan få adgang til de bedste GenAI værktøjer og de mindre ressourcestærke organisationer uden samme adgang, som kan skævvride graden af indflydelse uproportionalt.

11.3 Fremtidig forskning

Fremtidig forskning kan tage to retninger, svarende til de to måder, hvorpå AI-redskabet kan forbedres som diskuteret i implikationsafsnittet.

11.3.1 Forskning i intern forbedring

Fremtidig forskning kunne undersøge, om GenAI kan blive bedre til at klassificere dokumenter specifikt i en dansk kontekst. Sørensen har fremhævet at der er potentialet ved at opdele prompts yderligere og fine-tune sprogmodellerne i fremtidige GenAI værktøjer baseret på detaljeret dansk data, som fremtidig forskning kan fokusere på (bilag 28). Hvis det viser sig, at GenAI *ikke* kan forbedre dokumentklassifikation i en dansk politisk kontekst, vil det give en klar indikation af begrænsningerne ved at anvende GenAI til dokumentklassifikation inden for public affairs i Danmark.

Denne retning af forskning er vigtig, da den adresserer den praktiske anvendelighed og tilpasning af GenAI og sprogmodeller til lokale behov, som allerede er blevet identificeret som en potentiel udfordring, og ville kunne belyse potentialet ved at benytte GenAI i public affairs som en katalytisk ressource.

11.3.2 Forskning i eksterne forbedringer

En anden potentiel retning er at undersøge muligheden for at udvikle en RAG-model som en form for databaseværktøj (Merritt, 2023). SJ&K har udtrykt interesse i at se, hvordan en sådan model kan forbedre deres evne til at hente og analysere relevante dokumenter effektivt (SJ&K1). Dette vil kræve en omfattende evaluering af en RAG-models præstationsmålinger, herunder præcision, sensitivitet, og den praktiske anvendelighed i virkelige scenarier. Zero Carbon Shipping har også udtrykt, at en RAG-model kunne være værdifuld for at skabe mere dybdegående og relevante analyser, som vil kunne gavne public affairs organisationer med et hurtigt overblik over politiske udviklinger (ZCS2). Ved at forske i brugen af RAG-modeller, vil man kunne belyse hvordan GenAI potentielt kan skabe værdi for public affairs organisationer ved en anden anvendelse af GenAI end det AI-redskabet benytter. En sådan tilgang kan give indsigt i, hvordan GenAI kan forbedre arbejdsprocesser og effektivisere informationsindsamling, hvilket kan være afgørende for at udøve bedre public affairs.