



Artificial Intelligence og litteratursøgning i biblioteksregi

I de sidste to år, har Artificial Intelligence (AI) været på dagsordenen hos Det Kgl. Bibliotek (KB), hvor Københavns Universitetsbibliotek (KUB) og Aarhus Universitetsbibliotek (AUL) sammen har testet søgesystemer, der fungerer ved hjælp af AI.

Kan AI søgesystemer anvendes i en litteratur- og kildesøgnings kontekst i forskning og studier? Særligt med henblik på systematisk søgning og søgemetoder hvor gennemsigtighed, reproducerbarhed, dokumentation og pålidelighed er nøglen.

AI er et felt inden for informationsvidenskab, som er i hastig udvikling. Dette er særligt vigtigt i forhold til den voksende mængde artikler og data, der dagligt bliver større. AI har potentialet til at analysere store mængder tekster og via matematiske formler, at eftergøre menneskers måde at lære,

træffe beslutninger og løse problemer på. Opgaven kan oftest udføres med en hastighed, der langt overstiger et menneskes evne.

Om vores projekt

Projektet har udspiring i Københavns Universitetsbiblioteks innovationsstudietur til Helsinki 2019, hvor AI-erfaringer fra University of Helsinki blev belyst. Med baggrund i interessen for AI blev en projektgruppe søsat i 2020, hvor det blev besluttet hvilke AI-systemer, der skulle udvælges til test.

I projektets første leverance, identificerede vi 16 AI-systemer, hvoraf to, Iris.ai og Yewno.discover, levede op til vores specifikationskrav:

- Bruge en eller flere AI-drevne algoritmer
- Være designet til at understøtte akademisk litteratursøgning

- Være designet til at understøtte discovery af relateret litteratur/koncepter
- Være tilgængelig for test over en 2-3 årig periode
- Være egnet til anvendelse i fuldttekstressourcer, databaser med referencer og abstracts inden for de sundhedsvidenskabelige discipliner
- Have klare politikker og tilladelser vedrørende dataindsamlingsadfærd og privatlivspraksis
- Understøtte ansvarlig udførelse af forskning, såsom at levere funktionaliteter, der understøtter metodisk gennemsigtighed, reproducerbarhed og god citeringspraksis, herunder dokumentation af de anvendte søge- og rangerings-/klyngealgoritmer.



Metode og tests

Projektets metode er bygget op i tre faser:

- Udvælgelse af AI-systemer på baggrund af kravspecifikationer (Leverage¹)
- Tænke-højt tests af de to udvalgte systemer for at undersøge funktion, effektivitet, transparens, dokumentation og værdi for slutbrugerne.
- Hackathon med forskere og informationsspecialister med efterfølgende kvalitetsvurdering af resultaterne.

Tænke-højt tests giver viden om og forståelse af, hvordan brugerne opfatter systemet, dets funktionalitet samt systemets fejl, mangler og potentialer.

Testpersonerne (informationsspecialister) kommenterede hvordan systemet kunne bruges på universitetsbiblioteket.

Vores Hackathon var bygget på observationer af brugeradfærd, feedback fra testerne (informationsspecialister og forskere), en bibliometrisk analyse af de fundne publikationer og en ekspertvurdering af publikationernes kvalitet. Testpersonerne valgte deres egne tilgange til AI-systemerne for at løse en fælles case.

Vi havde en kontrolgruppe, som søgte i velkendte databaser såsom PubMed, Web of Science og Google Scholar, for at undersøge variation i søgeprocessen, resultaterne og oplevelser sammenlignet med AI-systemerne.

Udvalgte observationer

AI-systemerne tilbyder noget nyt i deres grafiske interface, et aspekt der blev vægtet højt af testpersonerne. Disse interfaces inviterer til at gå på opdagelse i annoterede node-graphs, der visuelt synliggør relationer mellem begreber og hvordan begreberne anvendes i dokumenterne.

Samtlige testpersoner farede dog vild i kryds-referencer og links mellem begreber, der i sidste ende forvirrede mere end de gavned. Testpersonerne kæmpede ligeledes med at dokumentere søgningen og forstå hvad der skete "bag facaden".

Særligt informationsspecialisterne var ikke imponerede over udfaldet af de søgninger, de foretog. Dette kan skyldes, at AI-systemerne søgningsmæssigt

er begrænset i forhold til de traditionelle bibliografiske databaser, som ofte bygger på en "klik og play" søgeform, og dermed bliver informationspecialisternes professionelle faglighed og viden om systemarkitektur udfordret.

Kvaliteten af de fremfundne publikationer blev vurderet af ekspertgruppen, der udtrykte lav tillid til publikationernes robuste videnskabelige grundlag. Derfor, for informationspecialisten, kræver AI-drevet søgning, ud over søgefaglig viden, en systemforståelse samt et disciplinfagligt kendskab i relevansvurderingen.

Konklusioner

Resultaterne viser at Iris.ai og Yewno.discover kan bruges i opstartsfasen til opgaveskrivning, da de er egnede til at udforske og undersøge begreber, finde tværdisciplinære emner og dermed generere ideer og hypoteser.

Men på nuværende tidspunkt konkluderer vi, at Iris.ai og Yewno.discover ikke er særlig effektive.

De to AI-systemer fungerede som "meta-søgemaskiner", og testpersonerne kunne se værdien i at søge ned i en stor mængde litteratur gennem én grænseflade. Systemerne kan dermed supplere de databaser som universitetsbibliotekerne udbyder, særligt hvis man betaler for at indlejre systemerne i bibliotekskatalogen.

Begge AI-systemer fremhæver Open Access (OA) kilder, og dermed supplerer OA strategier på universiteterne og nationalt. Der er dog også udfordringer med systemernes begrænsede funktionalitet, som kan udfordre slutbrugers ambition med en søgning bl.a. i forhold til antal begreber, der kan søges på én gang, høje krav til disciplinviden for at kunne udføre en hen-

sigtsmæssig søgning, og manglende transparens i -og dokumentation af søgeprocessen.

AI og litteratursøgning på KB - what's next?

De testede AI-systemer er på nuværende tidspunkt for umodne til at udføre pålidelige akademiske søgninger. Men AI-systemerne hjælper os til at forstå sammenhængen mellem begreber og udforske begreber på en holistisk måde frem for den lineære måde vi som informationspecialister er vant til. Det vil give en ret grundig forståelse af det søgte begreb.

Derfor vil vi fortsat gerne monitorere udvikling af AI og informationsøgning, og lære mere om hvorvidt de understøtter forskningsintegritet og hvordan informationskompetencer ligeledes burde udvikles.

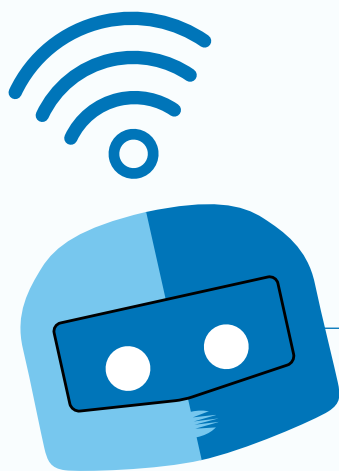
¹Læs leverance 1 her: Lorna Wildgaard, Solveig Sandal Johnsen, & Julie Kiersgaard. (2020).

Delivery 1: Selection of AI software. Zenodo. <https://doi.org/10.5281/zenodo.4279009>

²Skabeloner for tænke-højt tests af Iris.ai (<https://zenodo.org/record/5350166>) og

Yewno.discover (<https://zenodo.org/record/5350631>)

³Hackathon protokollen kan læses her: <https://doi.org/10.5281/zenodo.5349887>



Hvad er en AI-drevet søgning?

AI-systemer i dette projekt bruger følgende byggeklodser i kombination eller hver for sig:

"Aggregated behaviour": Er en søgeinterface, der integrerer søgeresultater i forskellige formater og fra forskellige kilder i en enkelt resultatside. Søgealgoritmen lærer fra søgepersonens adfærd og "click throughs" for at fremhæve de mest relevante resultater.

"Recommendations": En algoritme beregner og vægter forskellige søgemodeller for at fremme de mest relevante resultater. Disse beregninger kan inkludere information om søgepersonens lokation, tidligere søgninger samt lignende søgninger og adfærd af andre søgepersoner.

"Semantic search": Når algoritmen leverer en domæne-specifik fortolkning af søgetermerne. Fokus er på hvad termerne betyder i den kontekst de optræder i, i de søgte materialer.

"Clustering and classification": Når algoritmen trænes til at præsentere termerne i kategorier og i semantiske grupperinger, som slutbrugeren kan dykke ned i for at øge søgningens specificitet.