



Hvordan kan KI effektivisere Informasjonsforvaltning?

En liten PoC med stort potensiale innen begrepsbruk og datakildekartlegging på bærekraftsområdet

Faglig arena for informasjonsforvaltning og datadeling

Brønnøysundregistrene avd. Forretningsutvikling og Arkitektum AS



Brønnøysundregistrene

08.05.2024

PoC deltagere

Digitaliseringsrådgiver/løsningsarkitekt og systemutviklere fra
[Arkitektum AS](#)



Rådgivere fra Brønnøysundregistrenes avdeling
Forretningsutvikling (FU) - Faggruppe Informasjonsforvaltning,
Bærekraft, Datadelingstjeneseter og Dataanalyse (KI)



The image shows the exterior of a modern building with a dark facade and large glass windows. The name 'Brønnøysundregistrene' is written in large, white, 3D letters on the building's facade. In the foreground, there is a balcony with several dark grey tables and benches. The background features a coastal town with houses, a harbor with boats, and a large body of water under a blue sky with light clouds. Mountains are visible in the distance.

Brønnøysundregistrene

Hvem er Brønnøysundregistrene?

- Vi har 17 av landets viktigste registre og flere andre tjenester
- Opplysningene fra registrene våre brukes på tvers i både offentlig og privat sektor og gjør dialogen med det offentlige effektiv, samordnet og trygg
- 3 lokasjoner i Norge
- 430 ansatte

Bærebjelke i det norske samfunnet

- Vi tilbyr en automatisert og digital registerforvaltning gjennom hele verdikjeden
 - 95 % digital innrapportering
 - 71 % helautomatisk saksbehandling
 - 99,99 % digital tilgjengeliggjøring – 107 oppslag i sekundet
 - Registrering i våre registre gir publisitet, notoritet og legitimitet
- Vi tilbyr datadelingstjenester som tilrettelegger for et datadrevet samfunnet
- Vi har informasjon om alle virksomheter; næringsdrivende, frivillig og offentlig sektor



Hvem er Arkitektum AS ?

- Etablert i 2008
- 19 ansatte lokalisert Bø i Telemark
- Bistår offentlig sektor innenfor områder som
 - digitaliseringsrådgivning
 - standardisering
 - informasjonsmodellering
 - løsnings-/system- /IT-arkitektur
 - prosjektledelse
 - systemutvikling
- Kunder
 - Direktoratet for Byggkvalitet
 - Kartverket
 - Kommunal- og distriktsdepartementet
 - Arbeidstilsynet
 - KS
 - Arkivverket
- Fra konsulenter til hele produktteam



Arkitektum Prosjektteam KI, Informasjonsforvaltning og bærekraft



Øystein Thoen Sisjord

Systemutvikler



Leif Halvor Sunde

Systemutvikler



Per Ola Stålberg

Systemutvikler



Tor Kjetil Nilsen

Løsningsarkitekt/Digitaliseringsrådgiver

Teknologi Bærekraft

= «Twin transition»



99 %
SMBer

700 mrd.
NOK



Vi vil gjøre det enkelt for alle virksomheter i Norge
å dokumentere bærekraft



Hvor står vi og hva gjør vi på bærekraftsområdet?

- Implementering av bærekraftsdirektivet i Norge
- Rollen til Brønnøysundregistrene
 - Innsamlingsmyndighet
 - Forenklingsrollen krever offentlig – privat samarbeid
- I prosess for å bygge en basisløsning for Regnskapsregistret – kort sikt
- Satsingsforslag for modernisering av Regnskapsregistret
- Datadelingstjeneste for å gi bedriftene enkelt tilgang på data om selg selv i det offentlige for å bruke mindre tid på datainnhenting og mer på omstilling
- Det er et behov for en omforent forståelse av begreper og informasjon om bærekraft
- Deltar i et nordisk nettverk for å dele erfaringer og kunnskap



Hvorfor denne PoCen?



Hvordan kan generativ KI hjelpe oss med å effektivisere og optimalisere informasjonsforvaltningen på bærekraftsområdet?

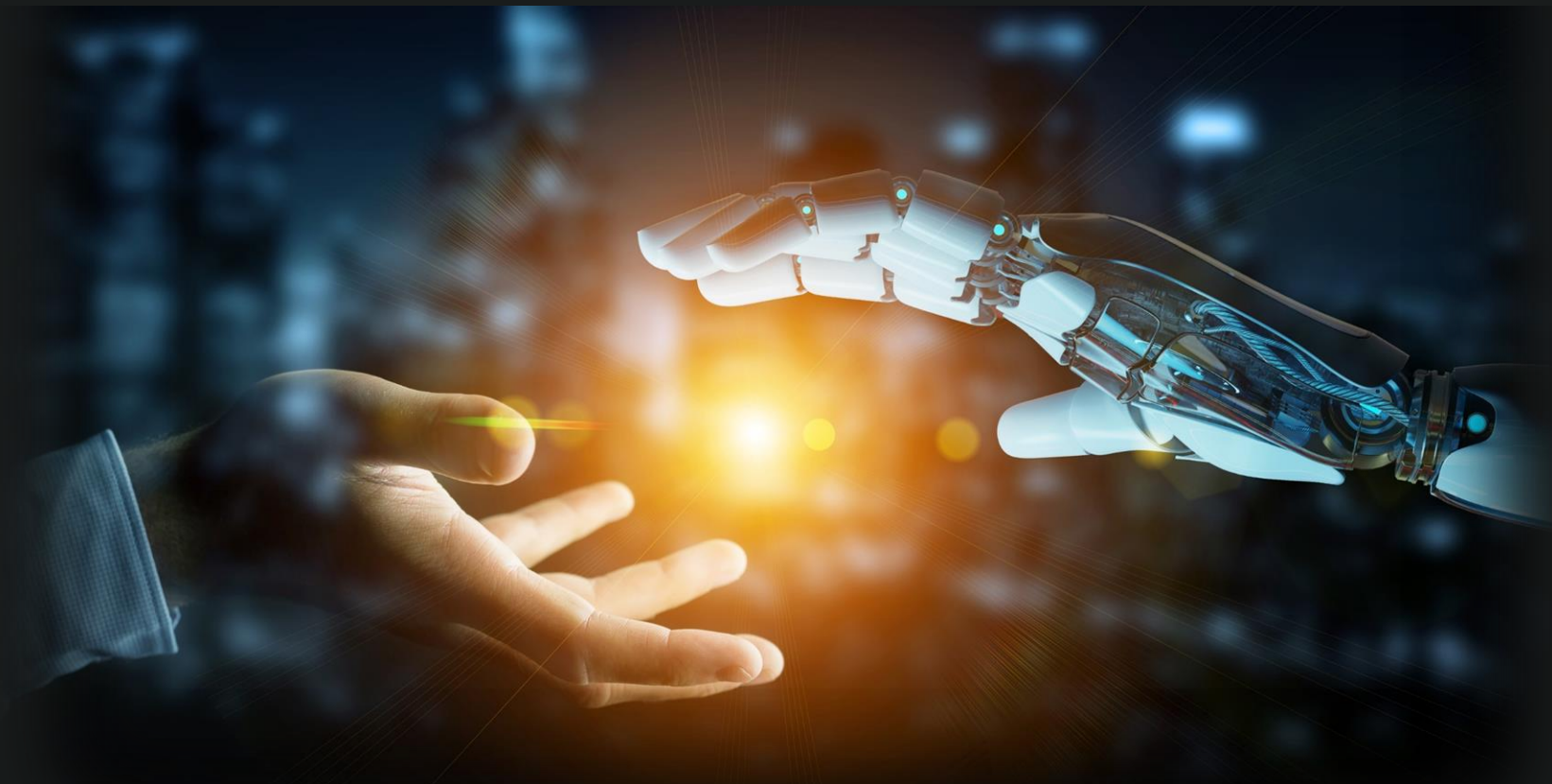
Kan vi få en assistent til å være ekspert på begrepsbeskrivelser og datakildekartlegging?

Bærekraftsfaktorer:

Environmental (Miljø)

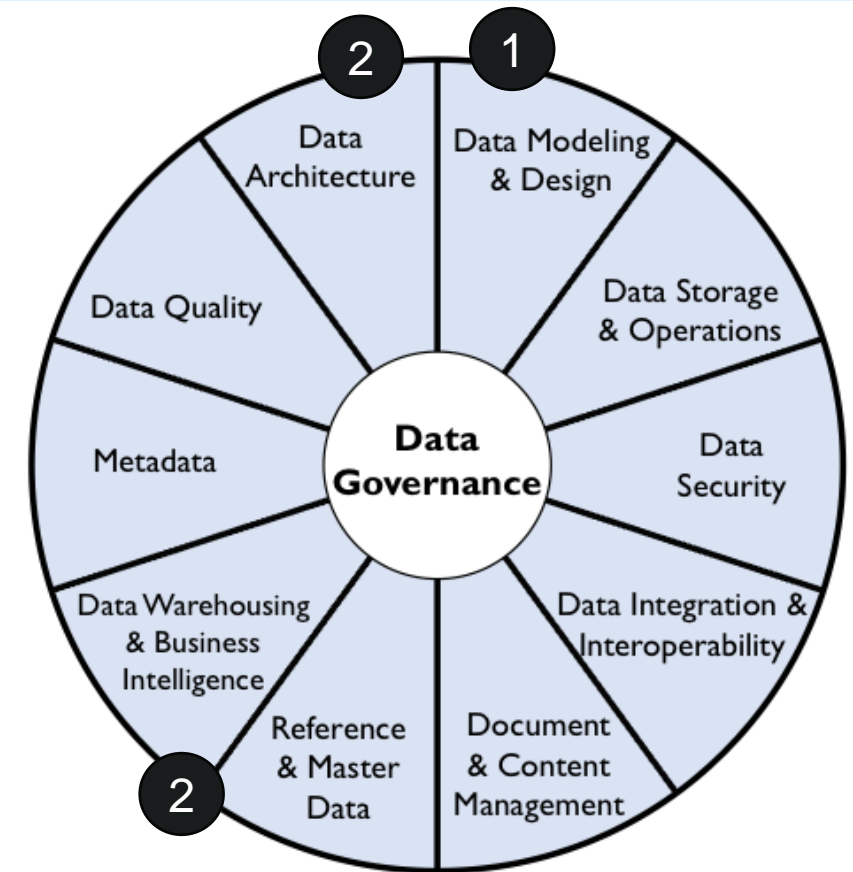
Social (Sosiale)

Governance (Forretningsetikk)



To Informasjonsforvaltningscase er valgt i denne PoC-en

- 1 Begrepsbeskrivelser** – kan en KI-assistent gi forslag til ESG-begrepsdefinisjoner som kan brukes i informasjonsmodeller? Resultater er i tråd med nasjonale standarder og BR rammeverk. Målgruppe: informasjonsforvaltere og SMB
- 2 Datakildekartlegging** – kan KI identifisere relevante kilder og datasett, og organisere i ulike ESG-kategorier? Målgruppe: informasjonsforvaltere



Copyright © 2017 DAMA International

dagens situasjon og ønsket effekt

Relevans

Begrepsdefinisjoner er en essensiell del av virksomhetsarkitektur for alle dataløsninger og -utvekslingstjenester (del av TOGAF).

Datakildekartlegging er et av de første stegene i alle datadelingstjenester.

Dagens situasjon

BR har for tiden kun manuelle prosesser for begge casene, som er relevante og kritiske for alle nye og oppdaterte interne datasystemer og -delingsløsninger.

Tidsbruk

Gjennomsnittlig tar 1 begrepsdefinisjon ca. 3.5 timer. BR har per i dag 850 begrep i begrepskatalogen.

En datakilde tar ca. 2 timer å identifisere ved manuelle søk. Disse estimatene er basert på BRs erfaring, og kommer an på mange faktorer, selvfølgelig!

Totalt har ESG ca. 1100 nye dataelementer, med én eller mange begrep, og kilder per element.

Ønsket effekt

Sluttbrukerne finner relevant og korrekt ESG informasjon raskt og effektivt.

BR:
- lavere kostnader ved bruk av KI
- ny innsikt: identifisere flere relevante datasett og mer informasjon enn ved manuelle søk

Demo 😊

Metodikk



Prosjektet har brukt kjente metoder som triple diamond med design thinking og prototyping i arbeidet.



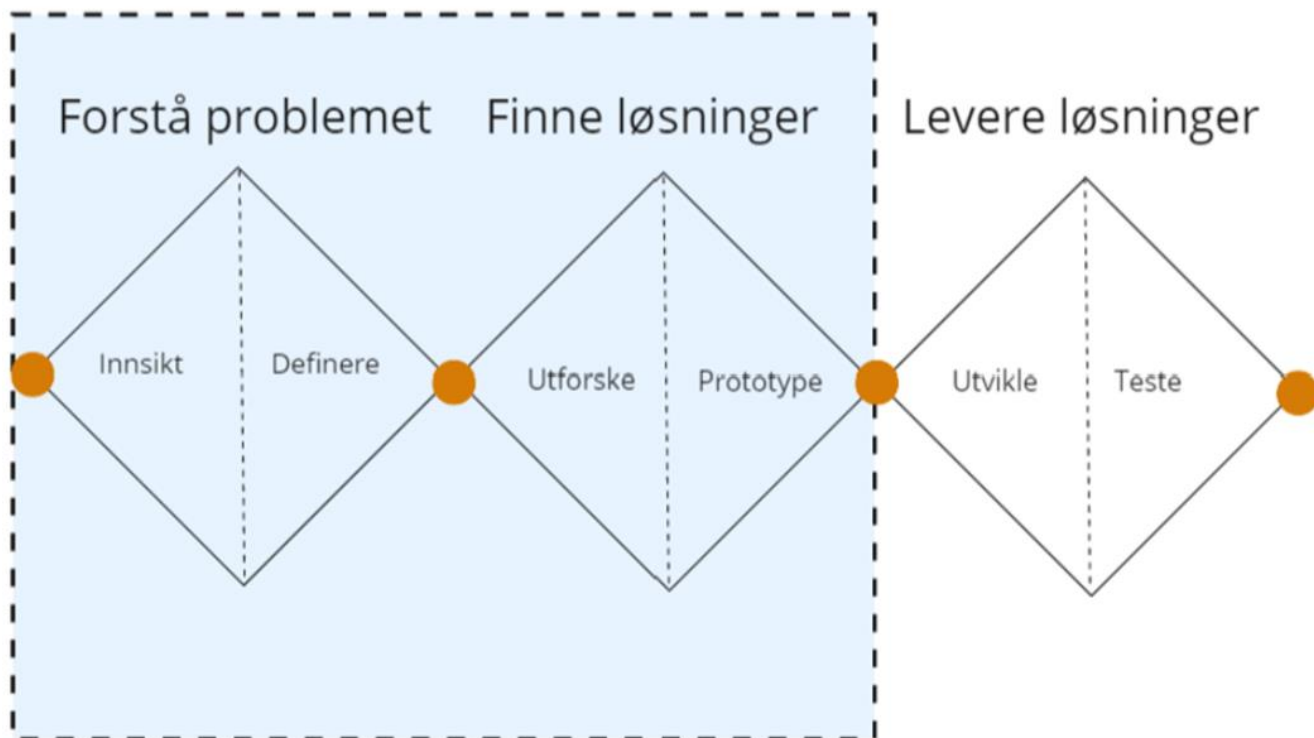
Deltakere fra Brønnøysundregistrene har bidratt med kunnskap innenfor informasjonsforvaltnings- og bærekraftsdomenet.



Prosjektet har sammen utarbeidet og definert problemstillingen, og utviklere fra Arkitektum har utforsket løsninger og utviklet prototyper basert på forskningen.



Prototyper er testet av hele prosjektgruppen.



Fremdriftsplan

Oppstart
januar



Begrepskartlegging

Rapport/presentasjon
og prototype 31.mars



Datakildekartlegging

Rapport

Budsjett 100 000,- + egeninnsats fra
Arkitektum og Brønnøysundregistrene

Case 1 Begrepsbeskrivelser



Brukerhistorie

- Som informasjonsforvalter så ønskes det bistand fra KI-assistent til å gi forslag på definisjoner til termer som brukes i informasjonsmodeller.
- Videre å teste ut om standarder som “Rammeverksdokument for begrepsarbeidet i Brønnøysundregistrene” og “Forvaltningsstandard for begrepsbeskrivelser (SKOS-AP-NO-Begrep) (Digdir)” kan brukes av KI-tjenesten som struktur og prioritering/ranking av definisjonsresultatene.



Kilde

God på norsk språk og rammeverk
for begrepsbeskrivelser?



Norsk språk og rammeverk for begrepsbeskrivelser

- Enkelt å sette krav til hvordan resultatet skulle presenteres jfr. Forvaltningsstandarden og Termlosen
- Henter inspirasjon fra relevante treff i datakatalogen
- GPT 4 er bedre enn GPT 3.5 på oversettelser til andre språk og målformer

Forvaltningsstandard for begrepsbeskrivelser
(SKOS-AP-NO-Begrep)

Digitaliseringsdirektoratet (Digdir) / The Norwegian Digitalisation Agency



Rammeverk

Rammeverksdokument for
begrepsarbeidet i
Brønnøysundregistrene



Presis og entydig definisjon er ikke alltid så lett å teste på, men erfaringen etter testing er at definisjonen ofte får kvaliteten som jeg ønsker, ...

Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere ?

emmissions to air

Valgfri kontekst ?

Send

SKOS-ap-no

Anbefalt term	utslipp til luft (nb) utslepp til luft (nn) emissions to air (en)
Definisjon	utslipp av gasser, partikler eller dråper til atmosfæren som kan skade helse, miljø eller bygninger (nb) utslepp av gassar, partiklar eller dropar til atmosfæren som kan skade helse, miljø eller bygningar (nn) release of gases, particles, or droplets into the atmosphere that can harm health, environment, or buildings (en)
Fagområde(r)	Luftforurensning (nb)
Kilde	Informasjon om luftforurensning fra ulike kilder som fabrikker, kraftverk og kjøretøy, og dens effekter på helse og miljø. (nb) Informasjon om luftforureining frå ulike kjelder som fabrikkar, kraftverk og køyretøy, og dens effektar på helse og miljø. (nn) Information on air pollution from various sources such as factories, power plants, and vehicles, and its effects on health and the environment. (en)
Kilde-URI	https://luftkvalitet.miljodirektoratet.no/artikkel/artikler/kilder-til-luftforurensning/

Hvordan gi KI assistent forståelse av
kontekst til begreper?



KI assistent og forståelse av kontekst

- Underliggende systemkontekst mottas av KI-modellen
 - Språk
 - SKOS-AP-NO-Begrep
 - BRs Begrepsrammeverk
 - Instruksjoner som inkluderer forklaringer av SKOS-AP-NO-Begreps-egenskaper
 - Eksempler på svar i ønsket form
- Ekstra kontekst kan angis av bruker
- Systemkontekst fylles med relevante treff fra Felles datakatalog
- Stor variasjon i hvor utfyllende datasettene i Felles datakatalog er beskrevet – KI kan sette sammen informasjon på en annen måte og bruker flere datakilder.



Bilde generert av **DALL-E**, gjennom Azure OpenAI Studio



Eksempel uten utfylt kontekst

Begrepsassistent

Datakildeassistent

Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere ?

mengde mikroplast generert

Valgfri kontekst ?

Send

SKOS-ap-no

Anbefalt term mengde mikroplast generert **(nb)**

Definisjon totalt volum av små plastpartikler mindre enn 5 millimeter i størrelse som er produsert eller frigjort til miljøet **(nb)**

Definisjon totalt volum av små plastpartiklar mindre enn 5 millimeter i storleik som er produsert eller frigjort til miljøet **(nn)**

Definisjon total volume of small plastic particles less than 5 millimeters in size that are produced or released into the environment **(en)**

Fagområde(r) Miljøovervåkning **(nb)**

Merknad Mikroplast kan genereres gjennom nedbryting av større plastavfall eller kan være primært produsert for bruk i produkter som kosmetikk. **(nb)**

Kilde Miljødirektoratets informasjon om mikroplast **(nb)**

Kilde-URI <https://www.miljodirektoratet.no/aktuelt/nyheter/2024/januar-2024/mikroplast-fra-bildekk-forureinar-oslofjorden/>



Eksempel med utfylt kontekst

Begrepsassistent

Datakildeassistent

Velg versjon av AI-modell



ChatGPT 3.5 Turbo



ChatGPT 4 Turbo

Begrepet du ønsker å definere [?](#)

mengde mikroplast generert

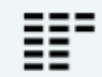
Valgfri kontekst [?](#)

bærekraft

Send

SKOS-ap-no

Anbefalt term	mengde mikroplast generert (nb)
Definisjon	totalt volum av små plastpartikler mindre enn 5 millimeter som produseres eller frigjøres til miljøet (nb)
	totalt volum av små plastpartiklar mindre enn 5 millimeter som vert produserte eller frigjevne til miljøet (nn)
	total volume of small plastic particles less than 5 millimeters produced or released into the environment (en)
Fagområde(r)	Miljøovervåkning (nb)
Merknad	Mikroplast kan oppstå ved nedbryting av større plastprodukter eller ved direkte utslipp fra industrielle prosesser og produkter. (nb)
Kilde	Inspirert av kontekst fra Felles Datakatalog og Bing-kontekst om mikroplast i miljøet. (nb)
Kilde-URI	https://data.norge.no/



Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere [?](#)

Valgfri kontekst [?](#)

SKOS-ap-no

Anbefalt term

as **(nb)**
as **(nn)**
as **(en)**

Definisjon

spansk sportsavis kjent for dekning av fotball, basketball og andre idretter, samt live hendelser og statistikk **(nb)**
spansk sportsavis kjend for dekning av fotball, basketball og andre idrettar, samt direktesendingar og statistikk **(nn)**
Spanish sports newspaper known for coverage of football, basketball, and other sports, as well as live events and statistics **(en)**

Fagområde(r)

Sport **(nb)**

Kilde

AS.com - Offisiell nettside for den spanske sportsavisen AS **(nb)**
AS.com - Offisiell nettstad for den spanske sportsavisa AS **(nn)**
AS.com - Official website of the Spanish sports newspaper AS **(en)**

Kilde-URI

<https://as.com>



Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere [?](#)

Valgfri kontekst [?](#)

SKOS-ap-no

Anbefalt term	aksjeselskap (nb) aksjeselskap (nn) joint-stock company (en)
Definisjon	juridisk person opprettet for å drive økonomisk aktivitet der eierne har begrenset ansvar for selskapets forpliktelser (nb) juridisk person oppretta for å drive økonomisk aktivitet der eigarane har avgrensa ansvar for selskapet sine forpliktingar (nn) legal entity established to conduct economic activity where owners have limited liability for the company's obligations (en)
Fagområde(r)	Næringsliv (nb)
Merknad	Et aksjeselskap er en selskapsform der eierne, kalt aksjonærer, har investert kapital i bytte mot aksjer, og deres ansvar for selskapets gjeld og forpliktelser er begrenset til aksjekapitalen. (nb) Eit aksjeselskap er ein selskapsform der eigarane, kalla aksjonærar, har investert kapital i byte mot aksjar, og deira ansvar for selskapet sin gjeld og forpliktingar er avgrensa til aksjekapitalen. (nn) A joint-stock company is a form of company where the owners, called shareholders, have invested capital in exchange for shares, and their liability for the company's debt and obligations is limited to the share capital. (en)
Kilde	Aksjeloven av 13. juni 1997 nr. 44 (nb)
Kilde-URI	https://lovdata.no/dokument/NL/lov/1997-06-13-44
Frarådd term	

Henvisning til kilder



Henvisning til kilder



404

Page not found

Oh Oh

**What
now?**

Kilde

- KI-modellers kunnskapsbase er begrenset.
- Henvisningen kan referere til utilgjengelig data.
- Vekting av kontekst mottatt via prompts og KI-modellers innebygde datagrunnlag.
- Opphavsrett

NYHET

**New York Times saksøker
OpenAI og Microsoft for
brudd på opphavsrett**

Hvordan kan dette forbedres?

- «Prompt-engineere» løsningen til å henvise riktig.
- Inkludere kilder i kontekst.
- Benytte teknikker utenfor KI for å identifisere kontekst.
- Benytte teknikker utover og sammen med KI for å identifisere kilder.



Bilde generert av **DALL-E**, gjennom Azure OpenAI Studio

- Utforske flere modeller: Perplexity, NorLLM, Gemini ...



Brønnøysundregistrene

Begrepsassistent

Datakildeassistent

Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere ?

fratreden

Valgfri kontekst ?

Send

SKOS-ap-no

Anbefalt term fratreden **(nb)**

Definisjon slutten på en persons ansettelsesforhold i en organisasjon, vanligvis frivillig **(nb)**

Fagområde(r) Arbeidsliv **(nb)**

Merknad Fratreden kan skje ved oppsigelse fra arbeidstaker eller arbeidsgiver, pensjonering eller andre frivillige avtaler om avslutning av ansettelsesforholdet. **(nb)**

Kilde <https://snl.no/fratreden> **(nb)**

Kilde-URI <https://www.youtube.com/watch?v=2hnhNUCEdX4>



Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere [?](#)

fratreden

Valgfri kontekst [?](#)

Send

SKOS-ap-no

Anbefalt term	fratreden (nb)
Definisjon	handling hvor en person avslutter sitt arbeidsforhold eller sitt offisielle verv (nb) handling der ein person avsluttar sitt arbeidsforhold eller sitt offisielle verv (nn) act where an individual terminates their employment or official position (en)
Fagområde(r)	Arbeidsrett (nb)
Merknad	Fratreden kan være frivillig som ved oppsigelse eller pensjonering, eller ufrivillig som ved avskjedigelse. (nb)
Kilde	Oversettelse av 'fratredelse' til engelsk i DinOrdbok, med setningseksempler. (nb)
Kilde-URI	https://www.dinordbok.no/norsk-engelsk/?q=fratredelse
Frarådd term	



Velg versjon av AI-modell

ChatGPT 3.5 Turbo ChatGPT 4 Turbo

Begrepet du ønsker å definere [?](#)

fratreden

Valgfri kontekst [?](#)

Lov om aldersgrenser for offentlige tjenstemenn

Send

SKOS-ap-no

Anbefalt term fratreden **(nb)**

Definisjon pliktig opphør av arbeidsforhold ved nådd alminnelig aldersgrense i henhold til lovgivning **(nb)**
pliktig opphøyr av arbeidsforhold ved nådd alminneleg aldersgrense i samsvar med lovverket **(nn)**
mandatory termination of employment upon reaching the general age limit according to legislation **(en)**

Fagområde(r) Arbeidsrett **(nb)**

Merknad I sammenheng med Lov om aldersgrenser for offentlige tjenstemenn, refererer fratreden til den lovpålagte avgangen fra en stilling når en ansatt når den alminnelige aldersgrensen. **(nb)**
I samanheng med Lov om aldersgrenser for offentlege tenestemenn, refererer fratreden til den lovpliktige avgangen frå ei stilling når ein tilsett når den alminnelege aldersgrensa. **(nn)**
In the context of the Law on Age Limits for Public Officials, resignation refers to the legally required departure from a position when an employee reaches the general age limit. **(en)**

Kilde Lov om aldersgrenser for offentlige tjenstemenn m.fl. § 2 fjerde ledd **(nb)**
Lov om aldersgrenser for offentlege tenestemenn m.fl. § 2 fjerde ledd **(nn)**
Law on Age Limits for Public Officials etc. § 2 fourth paragraph **(en)**

Kilde-URI <https://lovdata.no/dokument/NL/lov/1956-12-21-1>

Case 2 Datakildekartlegging



Brukerhistorie

- Som en SMB (De små og mellomstore bedriftene) vil jeg finne tilgjengelige datakilder for å genere vår ESG-rapport.
- Jeg vil gå ett sted for å innhente denne informasjonen, der alle datakilder er tilgjengelige i sanntid. Det samme er relevant for ESG-systemleverandører.
- Vi ønsker å forske på hvordan en KI-tjeneste kan gi oversikt over datakildene for å kartlegge relevant ESG-informasjon. Resultatene presenteres jfr. [Standard for beskrivelse av datasett, datatjenester og datakataloger \(DCAT-AP-NO\)](#).



Powerpoint stock photo

God oversikt over tilgjengelige datakilder på ett sted

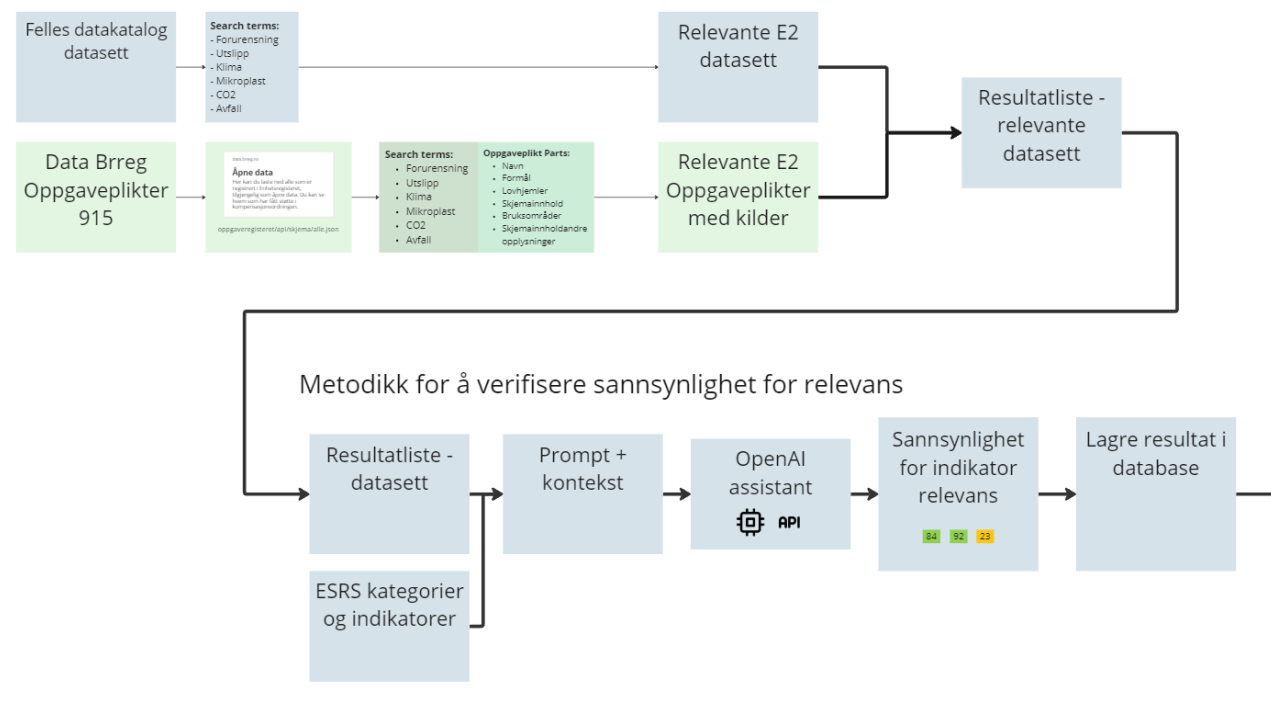
Begrepsassistent		Datakildeassistent			
Velg ESRS-kategori		Velg en indikator			
E2 - Forurensning		E2-4 §28b) - Mengde mikroplast generert			
Navn	Beskrivelse	Ansvarlig	Tilgang	Indikatorrelevans ^	Begrunnelse
Mikroplast i marine sedimenter	Mikroplast er små biter av alle typer plast som er mindre enn 5mm lange. Datasettet består av resultater av analyser av det øverste laget (0-2 cm) bunnsedimenter i norske hav og kystområder, som viser mengden av mikroplastpartikler i overflate-sedimenter. Mikroplast er målt i antall partikler i 1kg sediment tørrvekt.	Norges geologiske undersøkelse	PUBLIC	100 E2-4 §28b)	Datasettet viser mengden av mikroplastpartikler i marine sedimenter, og er derfor svært relevant for å måle mengde
Menneskelig påvirkning - Tapte fiskeredskaper (langs videotransekt)	Datasettet er en del av kartleggingsprosjektet Marine Grunnkart i Kystsonen og viser antall observerte tapte fiskeredskaper på havbunnen, inkludert tauverk, teiner og deler av ødelagt fiskeredskap. Datasettet er basert på videoanalyse og dekker enkelte kystområder der kartleggingsprosjektet Marine Grunnkart i Kystsonen har samlet inn videomateriale.	Havforskningsinstituttet	PUBLIC	90 E2-4 §28b)	Datasettet viser antall observerte tapte fiskeredskaper på havbunnen, inkludert tauverk, teiner og deler av ødelagt fiskeredskap.
Menneskelig påvirkning - Marint søppel per videotransekt	Datasettet er en del av prosjektet «Marine Grunnkart i Kystsonen», og viser antall observerte objekter (større enn 5 cm) på havbunnen som er kategorisert som søppel. Under søppel inngår alle elementer som er menneskeskapte, ikke inkludert fiskeredskap. Datasettet er basert på videoanalyse og dekker enkelte kystområder der kartleggingsprosjektet Marine Grunnkart i	Havforskningsinstituttet	PUBLIC	90 E2-4 §28b)	Datasettet viser antall observerte objekter på havbunnen som er kategorisert som søppel, og har dermed direkte relevans for
Søknad om tilskudd til spesielle miljøtiltak i jordbruket (SMIL)	Å fremme natur- og kulturminneverdiene i jordbrukets kulturlandskap og redusere forurensningen fra jordbruket, utover det som kan forventes gjennom vanlig jordbruksdrift. Prosjektene og tiltakene skal prioriteres ut fra lokale målsetninger og strategier.	LANDBRUKSDIREKTORATET	Ukjent	90 E2-4 §28b)	Datasettet omhandler søknad om tilskudd til spesielle miljøtiltak i jordbruket. Det kan være relevant for å vurdere mengde mikroplast
Kvotepliktige virksomheter: skjemaer for tillatelse og årlig rapportering	Søknadsskjema for tillatelse til kvotepliktige utslipp og ved søknad/melding om endring av tillatelsen. Skjema for årlig rapportering av kvotepliktige utslipp som skal via verifikatør og sendes inn med verifikasjonsrapporten som vedlegg til skjemaet. Kvotepliktige virksomheter skal bruke et søknadsskjema for tillatelse til kvotepliktige utslipp. Skjemaet skal også brukes ved søknad /melding	MILJØDIREKTORATET	Ukjent	90 E2-4 §28b)	Datasettet omfatter data mottatt på søknadsskjema for tillatelse til kvotepliktige utslipp og ved endring av tillatelse. Dersom
Årlig rapportering for avløpsanlegg	Den som er ansvarlig for et avløpsanlegg plikter årlig å rapportere nødvendige opplysninger om sine avløpsanlegg og utslipp fra disse til staten. Dataene er grunnlaget for å føre tilsyn med anleggene, publisere avløpsdata for offentligheten, rapportere avløpsdata til EU og utarbeide statistikk hos SSB.	MILJØDIREKTORATET	Ukjent	90 E2-4 §28b)	Datasettet omfatter årlig rapportering for avløpsanlegg, som er grunnlaget for å føre tilsyn med anleggene, publisere

For mye data til prosessering



KI og prosessering av data

- Datakildeassistenten har i oppgave å vurdere ESG-indikatorer opp i mot datasett fra Felles datakatalog og rapporteringsplikter fra Brønnøysundregistrenes Oppgaveregister.
- Prosessering av den totale mengden datasett og oppgaveplikter opp i mot ESG-indikatorene, resulterte i lang behandlingstid for KI-modellene.
- Prosjektet har benyttet GPT-3.5 for prosesseringen da den er vesentlig raskere enn andre GPT-modeller.



KI og prosessering av data

- Brukeropplevelsen til Datakildeassistenten var ikke god med den lange behandlingstiden. Derfor er hovedoppgaven til Datakildeassistenten flyttet bort fra brukeropplevelsen og lagret i en database.
- Videre forskningspunkter for å forbedre ytelsen:
 - Redusere antall oppgaveplikter og datasett som sjekkes
 - Teste andre modeller
 - Oppgradere tilgjengelige skyresurser



Powerpoint stock photo

Vurdere grad av relevans mellom ESG-indikator og datakilder



Relevans mellom ESG-indikatorer og datakilder

Begrepsassistent Datakildeassistent

Velg ESRS-kategori

E2 - Forurensning

Velg en indikator

E2-4 §28b) - Mengde mikroplast generert

Velg ...

- E2-4 - Forurensning av luft, vann og jord
- E2-4 §28b) - Mengde mikroplast generert
- E2-5 §34 - Mengde bekymringsfulle stoffer som forlater bedriften som utslipp, en tjeneste eller del av produkt/tjeneste
- E2-5 §34 - Total mengde bekymringsfulle stoffer som genereres eller brukes under produksjon eller som er anskaffet
- E2-6 - §40a) - Andel av nettoinntekter som kommer fra produkter eller tjenester som består av bekymringsfulle stoffer

Markfuktighet - DTW

Indikatorrelevans ^	Begrunnelse
100 E2-4 §28b)	Datasettet viser mengden av mikroplastpartikler i marine sedimenter, og er derfor svært relevant for å

Teknisk løsning

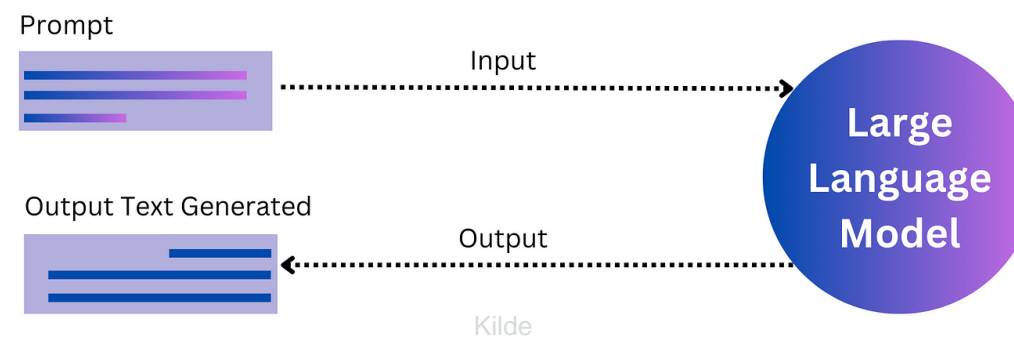


Intro til generativ KI

Generativ KI er kunstig intelligens som kan generere nytt innhold. For eksempel med en LLM.

Prompt er instruksjon brukeren gir KI-modellen. Ved å tolke instruksjonen, genererer KI-modellen nytt innhold til brukeren.

Prompt-engineering er kunsten å utforme forespørsler og instruksjoner for å oppnå ønsket respons eller best mulig svar.



Arkitektur

Består av 3 hoveddeler hvor alle har Azure som vert.

Nettsiden/Frontend

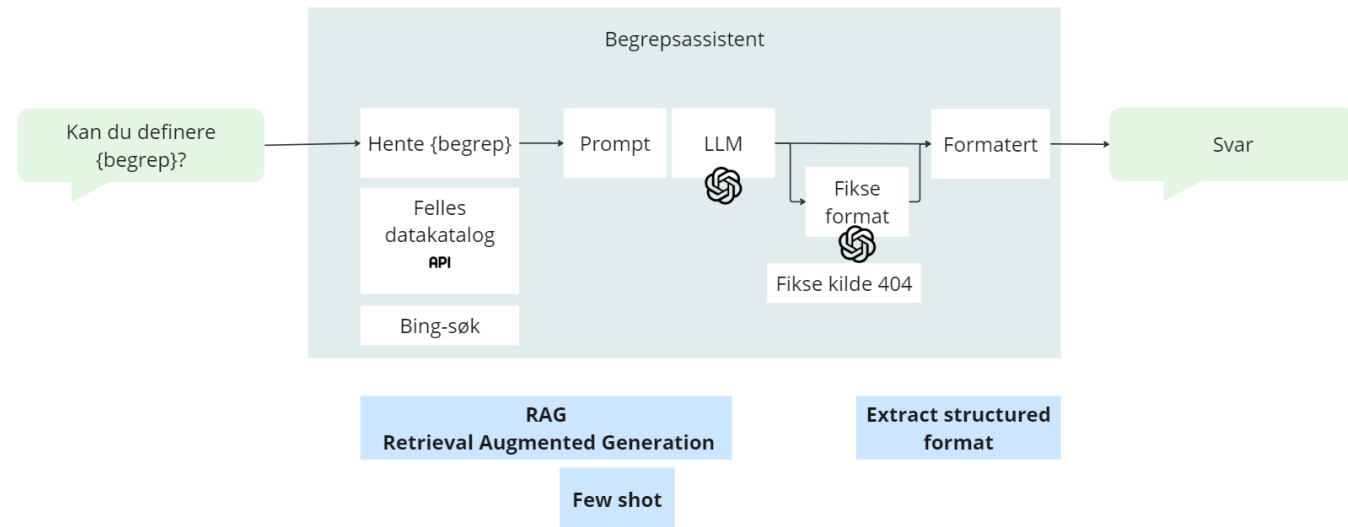
- Grafisk brukergrensesnitt
- React
- Kommuniserer med backend

Backend

- .NET8
- API-er til begreps- og datakildeassistenten
- Kommunikasjon med KI-modellen og nettsiden

Azure Open AI

- GPT-modeller
- Komponenter i Azure.
- Kalkulere respons basert på prompt og kontekst, og kommuniserer med backend.



[Kilde](#)

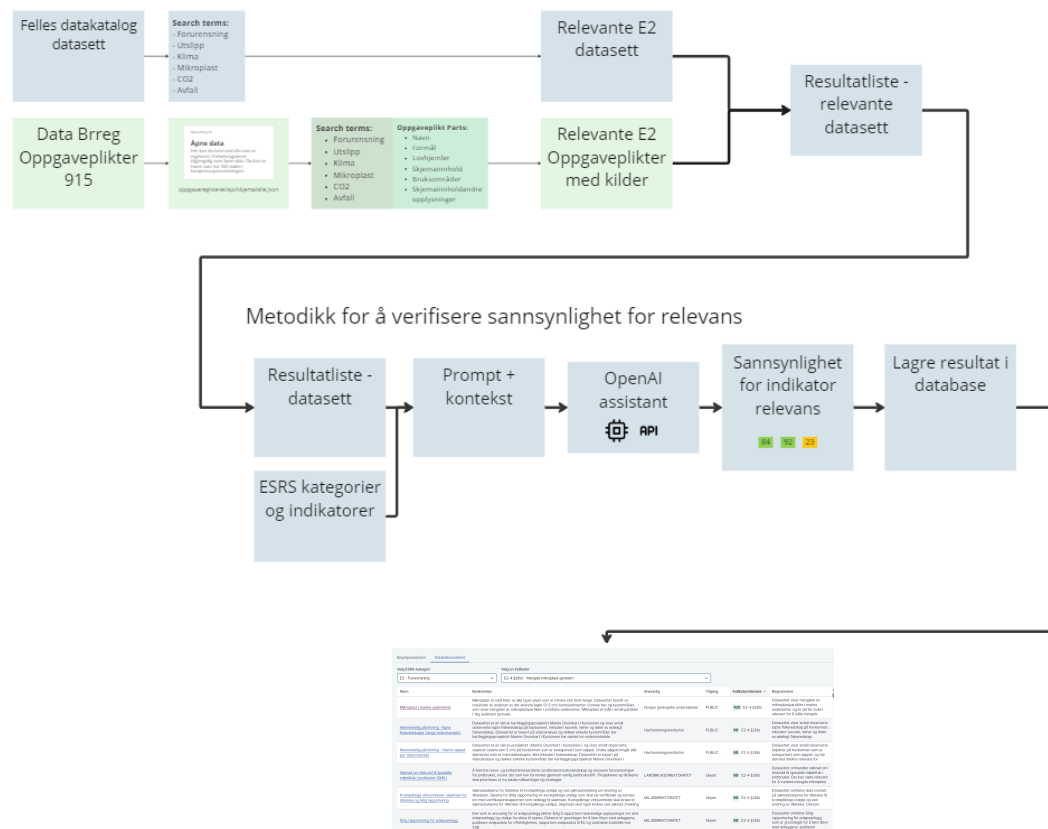
Løsning

Nettsiden er bygget opp hovedsakelig med komponenter fra designsystemet.no

Prompt-Engineering

- Few-shot
- Noen eksempler på spørsmål og forventet svar

Dersom KI-modellen besvarer oppgaven i feil form, er det benyttet en ny KI-modell til å rette feilene den første KI-modellen utførte automatisk.



Veien videre

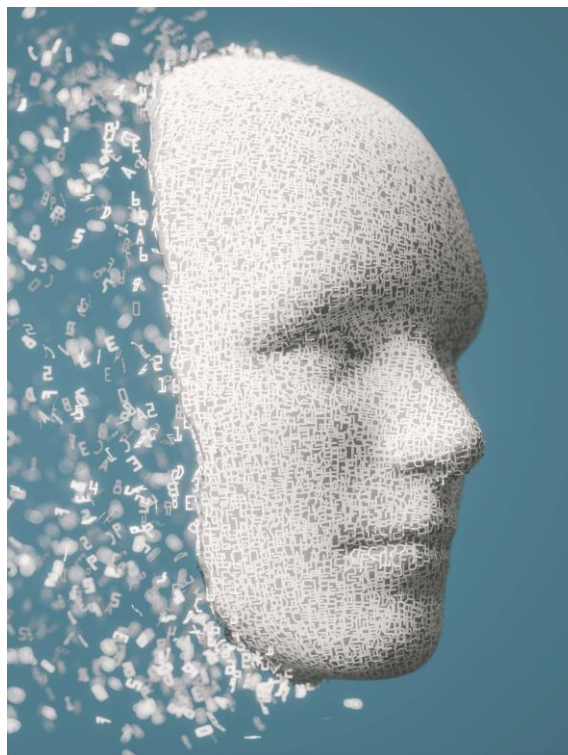


Usikkerhet og risiko - vurderinger

Denne PoCen – bruk av KI

- Åpne data – tilgjengelig på internett
- Amerikanske språkmodeller
- KI kreativitet/temperatur

Usikkerhet- og risikonivået ble vurdert som lavt



Videre utvikling og testing

- Kombinasjon av BR data og Åpne data
- Opphavsrett i forbindelse opptrening av KI språkmodeller - Open AI / Google opphavsproblematikk
- Nye norske språkmodeller
- KI i kombinasjon med andre teknologier – f.eks. RDF kunnskapsgraf, automatisert informasjonsmodellering

En risikovurdering før videre utvikling og testing bør gjennomføres på et tidlig stadium

Mange ønskede effekter i denne PoCen ble oppnådd

Her er et utvalg av positive effekter i prosjektet

- Bredde i informasjons-omfanget – KI kan sette sammen informasjon på en annen måte og bruker flere datakilder - eks. det ikke alltid god praksis på hvilken info som detaljeres i FDK. KI-assistenten bygger opp bedre bilde av tilgjengelig info (f.eks. setter sammen info fra FDK med info på internett).
- Effektivisering - konsistente svar, jobber raskt, finner mer informasjon raskere enn et menneske.
- KI-assistentene er flinke til å følge standarder og rammeverk – f.eks. SKOS-AP-NO-Begrep, BRs Begrepsrammeverk, fremstille i JSON.
- Potensiale for bedre kvalitet på begrepsarbeidet og datakartleggingen – men begge KI-assistentene trenger videreutvikling de kan tas i bruk.
- Amerikanske trente språkmodeller ble ikke en hindring for å få gode norske svar.
- Gode svar – temperaturen ble justert ned slik at KI-assistenten ikke ble for "kreativ" i sine svar.
- Oversettelser og målformer - Chat GPT 4 Turbo er god til å oversette til nynorsk og engelsk, som er en anbefaling i standarden for begrepsbeskrivelser.

Forbedringspotensialer og videreutvikling - Begrep

- Påliteligheten på kildehenvisninger i «begrepsassistenten» påliteligheten ikke god nok – må jobbes videre med
- Ranking / prioritering av kilder er ikke konsistent – her bør KI-assistentene trenes videre
- Hvordan kan en norsk språkmodell endre resultatene vi får? NorLLM lansert 15. mai 2024 – dette er interessant å se videre på



Forbedringspotensialer og videreutvikling – Datakildekartlegging

- Det ble forholdsvis lite tid til Datakildekartleggingen i PoCen. Ønsker å gå videre å utforske denne.
- Datakildeassistenten lister for mange resultater – søker for bredt per bærekraftsindikator. Vurdere om dette kan videreutvikles f.eks. i parallell med norsk versjon av bærekraftsindikatorerne.
- Hint fra nasjonale og globale statistikkilder er for omfattende i denne fasen, men et alternativ for neste trinn. Derimot ble informasjon om oppgaveplikter fra Oppgaveregisterets datasettet valgt.
- Alternativt kan Kunnskapsgraf (grafdatabase) brukes for å sjekke relasjoner og knytte datasett mot hverandre.



Neste steg for BR

- Presentere «Begrepsassistenten» som et potensiale for berikelse BRs begrepsprosess
- Utvide innsiktsarbeidet for «Datakildekartlegging»
 - Spesifikt mot bærekraftsinformasjon (ESG)
 - Nye case, som f.eks. kartlegging av Oppgaveregisteret
- Utforske bruk av KI i kombinasjon med RDF kunnskapsgraf

Takk for oppmerksomheten!

Ønsker du mer informasjon? Kontakt

Sissel Hovstein Olsen- sols@brreg.no

Tor Kjetil Nilsen – tor@arkitektum.no

