

Skatteetaten

AI-støttet arkivering & journalføring

Utvikling av prototype for automatisert håndtering av arkiv- og journalpliktig dokumentasjon i Skatteetaten

Fagleg arena for datadeling og informasjonsforvaltning

v/ fagdirektør for dokumentasjonsforvaltning Marcus Liebold

Oslo, den 12.02.2025

Agenda

- Bakgrunn / problem
- Behov
- Prosessen hittil
- Løsningsdesign
- Før-kommersiell anskaffelse
- Fase 1 – Krav til prototypen
- Fase 1 – Utvikling av prototyper / funn
- Fase 2 – Krav til prototypen
- Kommunikasjonsstrategi
- Tilgjengeliggjøring / implementering
- Hva betyr tiltaket for Skatteetaten?



Bakgrunn / problem:

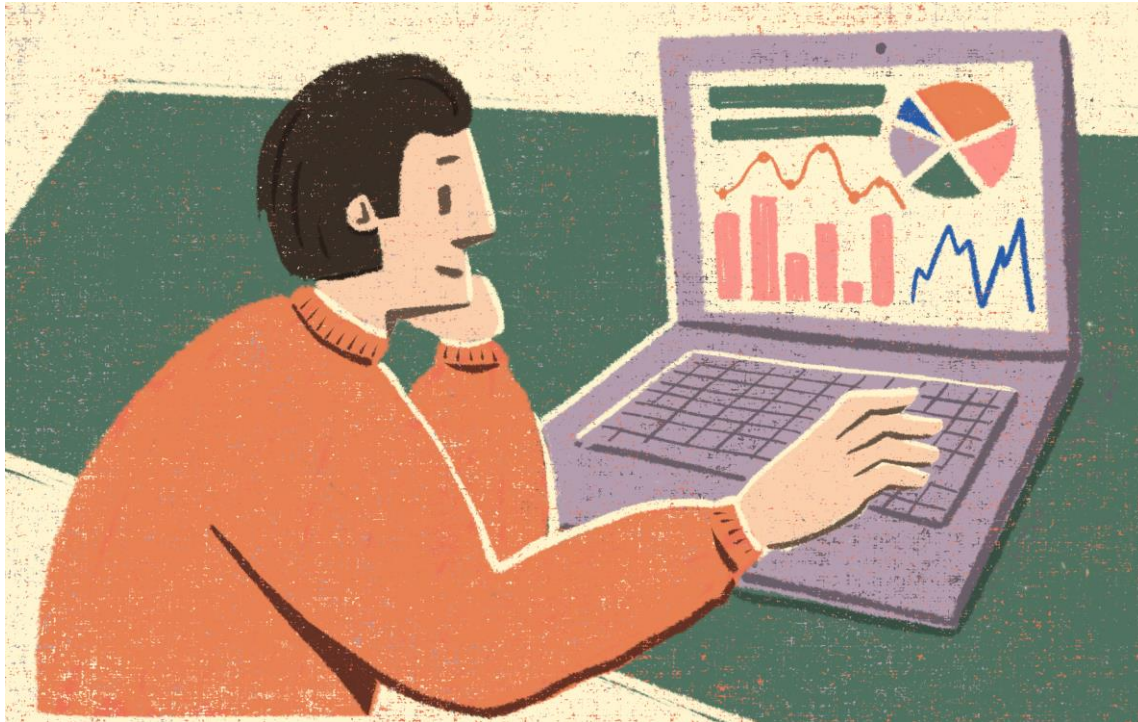
- Arkivloven § 6, arkivforskriften §§ 9 og 11
- Betydelige **utfordringer** i Skatteetaten knyttet til **forsvarlig arkivering** av arkiv- og journalpliktig dokumentasjon fra **e-postklienter** og **digitale samhandlingsløsninger**
- Årsaken til dette er at arkivfaglige prinsipper legger opp til at den enkelte ansatte har **selvstendig ansvar** for arkivering, **uten** at den nødvendigvis råder over **tilstrekkelig kompetanse** til å kunne utføre forsvarlig arkivering
- **Arkivering** krever til dels skjønnsmessige vurderinger og oppleves derfor som en **kompleks** og **tidkrevende** oppgave, som hindrer den enkelte å kunne bruke tid på kjerneoppgaver
- Manglende arkivering og journalføring har imidlertid gjentatte ganger ført til **presseoppslag** og **tilsynssaker** som truer Skatteetatens omdømme i offentligheten

Skattemyndighetene brøt lov i hemmeligholdt skattesak

Skattedirektoratet ønsket å skjerpe skatten på gunstige aksjeavtaler. Så bråsnudde de. Ingen av sakens dokumenter ble arkivert.

Publisert: 14. januar





Behov:

- Ivareta regulatoriske krav innen arkivering, personvern og informasjonssikkerhet
- Effektivisere Skatteetatens arbeidsprosesser knyttet til arkivering og journalføring
- Frigjøre kapasitet i linjeorganisasjonen
- Redusere risiko for alvorlige avvik

Systemstøtte:

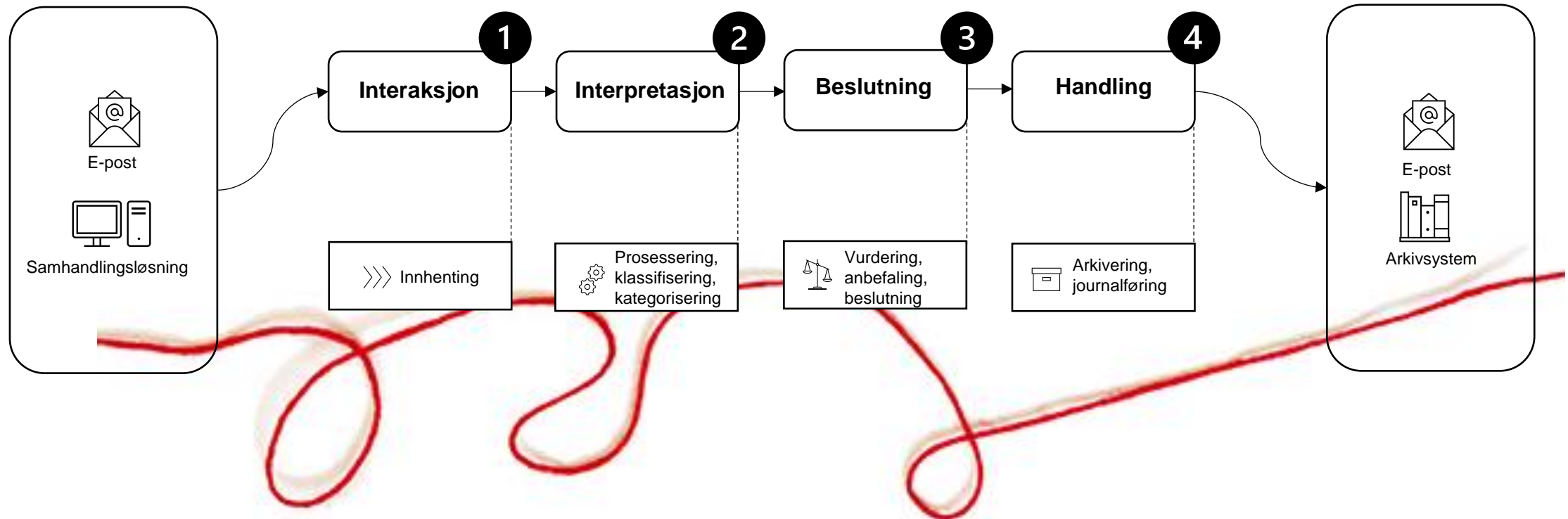
- AI / maskinlæringsløsninger som kan:
 - Prosessere tekst i ulike formater,
 - Analysere og gjenkjenne mønstre i tekstene,
 - Kategorisere innholdet i tekstene
 - Vurdere om informasjonen er av privat karakter, om den er journalpliktig og/eller arkivpliktig
 - Presentere vurdering ovenfor den enkelte ansatte som så skal godkjenne resultatet av vurderingen



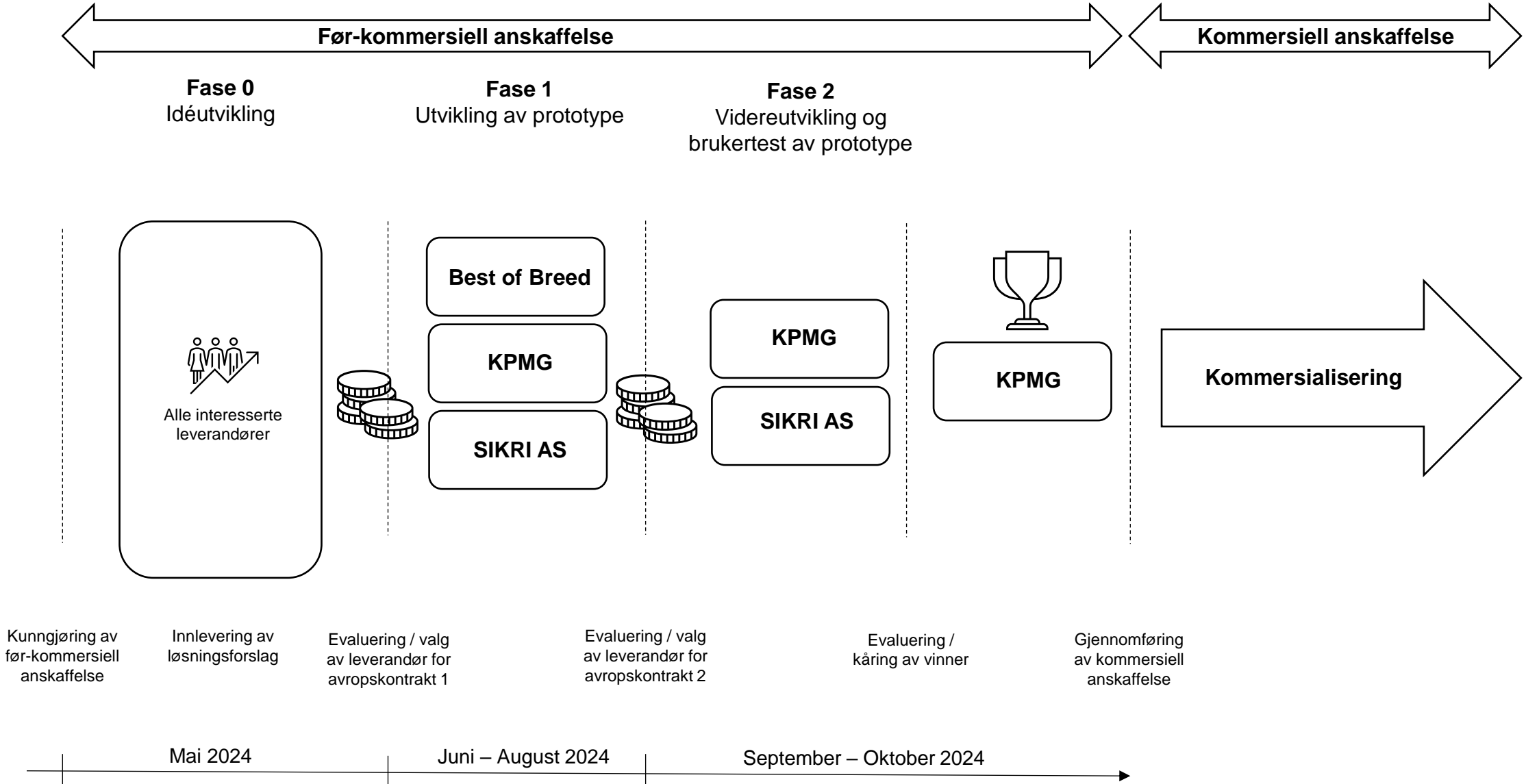
Proessen hittil:

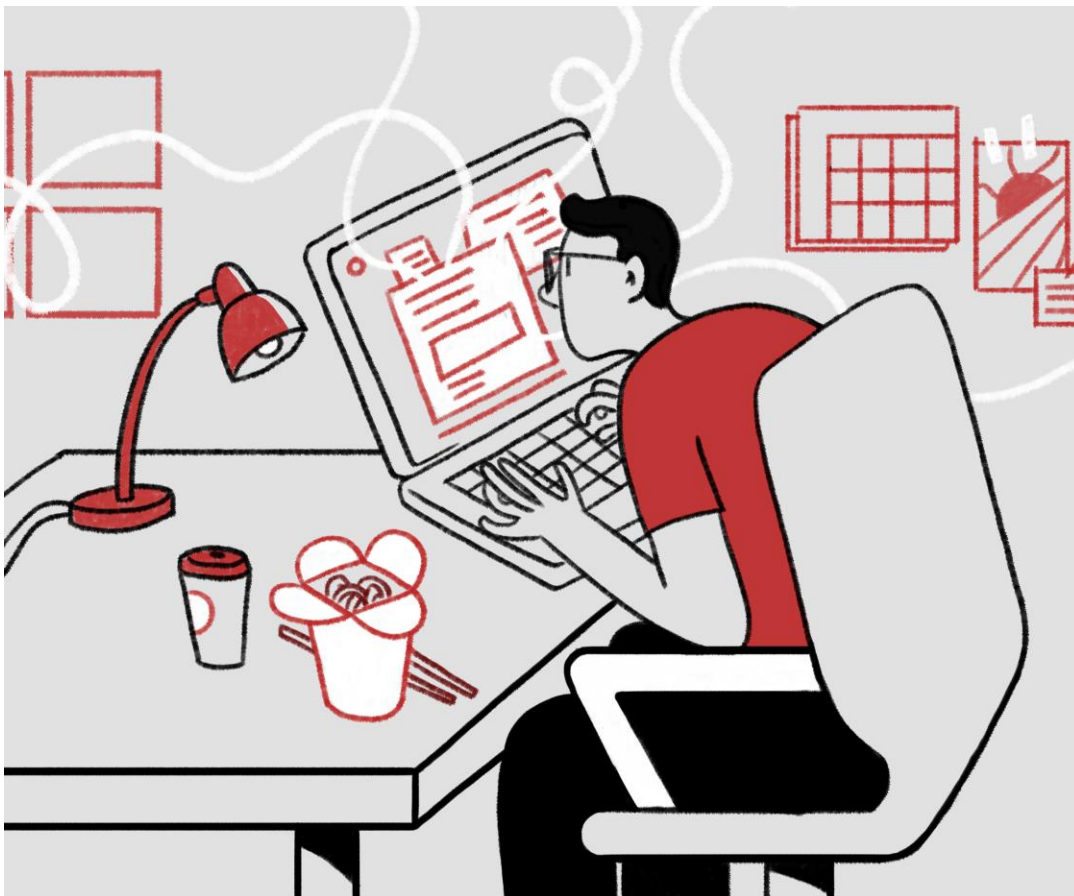
- Oktober 2023
 - Behovsmelding med problembeskrivelse & løsningsforslag med anbefaling til opprettelse av tverrfaglig arbeidsgruppe
- November 2023
 - Drøfting i tverrfaglig arbeidsgruppe
 - Casebeskrivelse for behandling i Skatteetatens eksperimenteringsløp
- Desember 2023
 - Utforming og gjennomføring av markedsundersøkelse (RFI)
- Mars 2024
 - Tildeling av 1,5 mill kroner fra Skatteetatens eksperimenteringsløp
- Mai – oktober 2024
 - Eksperimentgjennomføring – Utvikling av prototype for AI-støttet arkivering & journalføring

Høynivå løsningsdesign:



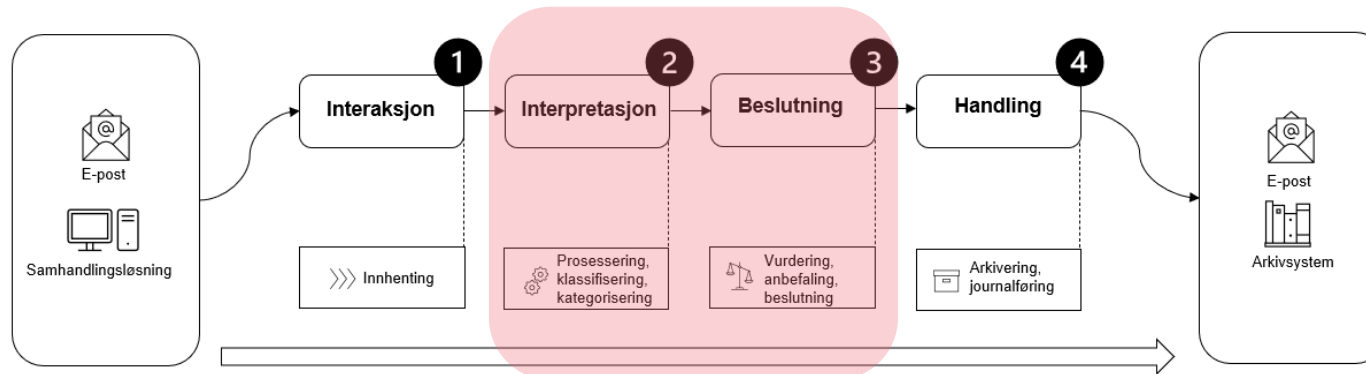
Metode for anskaffelse av forsknings- og utviklingstjenester:





Fase 1 – Krav til prototypen:

- Vurdere 100 anonymiserte testdokumenter ifht arkiv- og journalplikt
- Sluttresultatet skal være mulig å laste ned som en Excel-fil som viser:
 - Dokumentnavn
 - Arkivplikt / journalplikt (som to ulike kolonner)
 - Begrunnelse for hvorfor et dokument er vurdert som arkiv- og/eller journalpliktig (eller ingen av delene)
- Må fungere på Microsoft Azure (Skatteetatens sky-infrastruktur)
- Må kunne hostes på Skatteetatens egen tennant (klassifisering skal ikke gjøres utenfor Skatteetaten)
- Må ha grensesnitt som muliggjør overføring til arkivsystemet (Elements)
- Tillegg:
 - Vurdering på skjerming av innhold i dokumentet (tilgangsvurdering)





SIVILOMBUDET

SKATTEETATEN
Postboks 9200 Grønland
0134 OSLO

Vår referanse 2022/5128	Deres referanse 2022/3938	Vår saksbehandler Ellen Cathrine Beccer	Dato 31.01.2024
----------------------------	------------------------------	--	--------------------

Vedrørende ombudets uttalelse – Om Skatteetatens arkivering, journalføring og håndtering av etatsintern korrespondanse mv.

Det vises til Sivilombudets uttalelse 11. september 2023.

I uttalelsen kom ombudet blant annet til at det var uklart om Skatteetatens arkiv er i tråd med arkivlova § 6, og at det var uklart om journalene tilfredstilte de kravene som er oppstilt i arkivforskrifta § 10 første ledd bokstav c.

På denne bakgrunn ber ombudet om en orientering om hva Skatteetaten har foretatt seg etter ombudets uttalelse, herunder om svar på følgende spørsmål:

- Hva har Skatteetaten konkludert med vedrørende inndelingen av arkivet og journalføringen av korrespondanse innad i Skatteetaten? Har etaten vært i kontakt med Finansdepartementet og/eller Arkivverket i forbindelse med disse vurderingene?

Sivilombudet ber om svar innen tre uker.

Hanne Harlem
sivilombud

Dette brevet er godkjent elektronisk og har derfor ikke håndskrevet underskrift.

Avsender = eksternt forvaltningsorgan

Eksternt saksnummer = inngår i saksbehandling & har verdi som dokumentasjon

Brevtittel = viser til eksisterende saksbehandlingsløp

Skatteetatens saksnummer = eksisterer en sak i Elements admin

Konklusjon:

Dokumentet er journal- og arkivpliktig



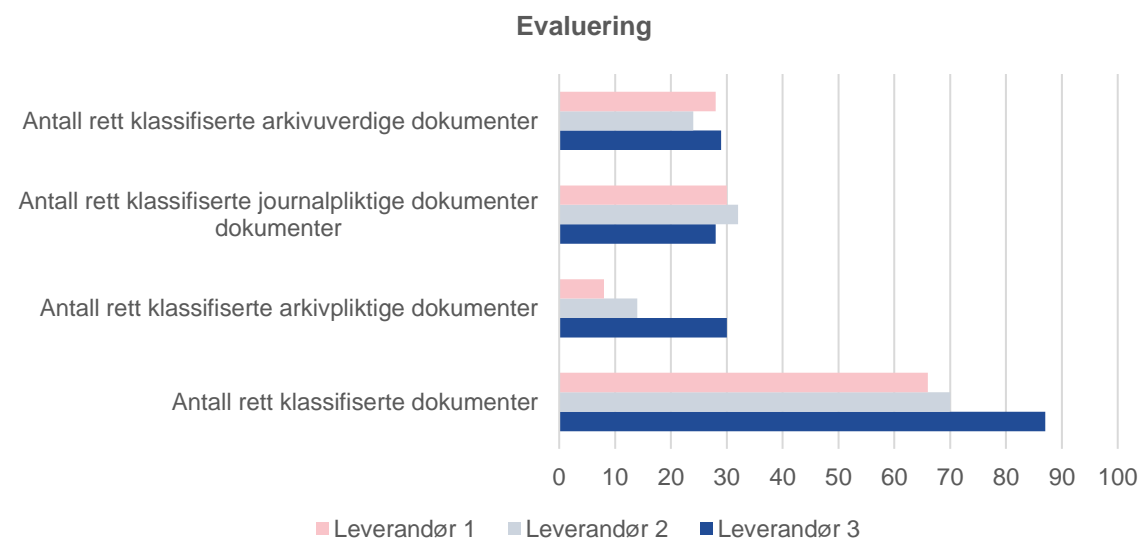
Produsent = eksternt aktør - Aftenposten

Innhold = innenriks / utenriks nyheter

Konklusjon:

Dokumentet er verken journal- eller arkivpliktig





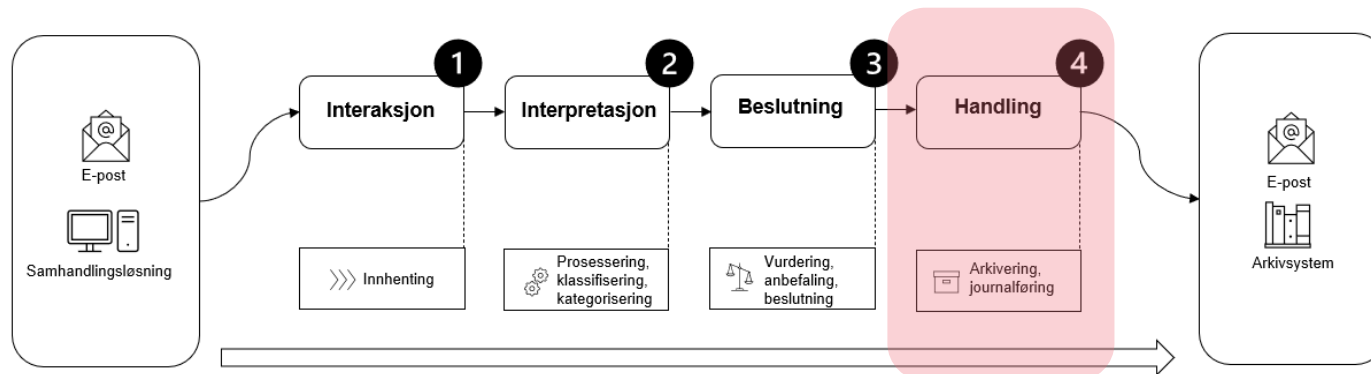
Fase 1 – Utvikling av prototype & funn:

- Gjennomført workshops med leverandører for avklaring av faglige spørsmål
- Sammenstilt & syntetisert 100 testdokumenter for klassifisering av arkiv- og journalplikt
- Utviklet et AI-basert verktøy for anonymisering av personsensitive opplysninger som kan benyttes til bl.a. besvarelse av innsynsforespørsler
- Gjennomført møter med leverandører for demonstrasjon av prototyper
- 2 av leverandørene har basert sin prototype på prompting til en GPT4-språkmodell, mens en leverandør har bygget egen språkmodell
- Evaluering av prototypene viser rett klassifisering på 66-87% av dokumentene
- Treffsikkerheten på klassifiseringen av journalpliktige dokumenter ligger på 85-96%
- Størst avvik ved klassifisering av arkivpliktige dokumenter (som forventet) – treffsikkerheten her ligger på 21-37% (en leverandør scorer 79%)
- Totale treffsikkerheten til språkmodellene er likevel høyere enn ved menneskelig vurdering



Fase 2 – Videreutvikling av prototypen:

- Konfigurere prototypen til å kunne ekstrahere nødvendige journalopplysninger fra journalpliktige dokumenter
- Generere JSON-fil med nødvendige journalopplysninger per dokument som skal brukes til å berike en matrise for en generisk journalrapport
- Konfigurere prototypen til å kunne generere standardiserte journalposttitler for journalpliktige dokumenter
 - med utgangspunkt i Skatteetatens fordelingsliste med skriveregler for enkelthenvendelser
- Utforme en detaljert beskrivelse av den ferdigutviklede prototypen som dokumenterer:
 - hvordan den overholder kravene i NOARK-standarden
 - hvordan den kan integreres med en NOARK-godkjent arkivløsning
 - hvordan overføringen av klassifiserte arkiv- og journalpliktige dokumenter utføres fra prototypen til en NOARK-godkjent arkivløsning
 - om den er tenkt å være en stand-alone-løsning eller en integrasjonsløsning og hvordan implementasjon av brukergrensesnittet er tenkt løst
- Evaluering av prototypenes funksjonalitet skjer på bakgrunn av et nytt testsett av dokumenter


Dokument

Fra: arkivplan23@gmail.com[arkivplan23@gmail.com]
 Sendt: 31.08.2024 00.06.17
 Til: SKD Innsyn OEP[SKD.InnsynOEP@skatteetaten.no]
 Tittel: Innsynskrav

Til: Skattedirektoratet
 Bestillingsdato: 31.08.2024

Svar på dette innsynskravet ønskes sendt med e-post til:
 arkivplan23@gmail.com
 Innsynskravet er bestilt av:

Dokumenter:
 Saksnr.: 2022/3938 | Dok nr. : 23 | Sekvensnr.: 28408/23
 Sak: Sivilombudet - Undersøkelse av eget tiltak - om divisjoner og regioner er selvstendige forvaltningsorganer i innsynssaker
 Dokument: Oversendelse av uttalelse - om Skatteetatens arkivering, journalføring og håndtering av etatsintern korrespondanse
 Dok.dato: 2023-09-11
 Journaldato: 2023-12-02
 Saksbehandler: M93878

Journalopplysninger

ID	Forklaring
1	Journalføringsdato (journaldato)
2	Saks- og dokumentnummer
2.1	Saksnummer
2.2	Dokumentnummer
3	Navnet til avsenderen eller mottakeren
3.1	Avsendernavn
3.2	Mottakernavn
4	Opplysninger om saken, innholdet eller emnet
4.1	Sakstittel
4.2	Dokumenttittel
5	Dateringen på dokumentet
5.1	Dokumentdato
6	Klasse etter organets ordning for klassifikasjon
6.1	Arkivkode (klassering)
7	Ekspedisjons- eller avskrivingsdato
7.1	Avskrivingsdato
8	Avskrivingsmåte (TE/BU)
9	Dokumenttype (I, U, N, X)
10	Tilgangskode og hjemmel
10.1	Tilgangskode
10.2	Hjemmel for skjerming av innhold

Datastruktur


Innhold: Krav om innsyn etter offentleglova JP1			
Sakstittel: Arkivering og journalføring i seksjon skatt			
DokType	Sak/dok nr:	Løpenr.:	Journaldato:
I	2023/734-2	24200/2024	02.09.2024
			Dok.dato: 31.08.2024
Tila.kode Hjemmel:		Klasseringa:	
		I01A01	
Avs./mottaker:	Navn	Avskr. dato:	Avskrivingsmåte:
Mottaker	Skattedirektoratet	03.09.2024	BU
Avsender	arkivplan23@gmail.com		
			Avskriv/besvar Inr.: 3

Journalrapport

From: Liebold, Marcus[Marcus.Liebold@skatteetaten.no]
Sent: 06.11.2023 11:09:56
To: Postmottak[postmottak@arkivverket.no]
Subject: Avlevering av Etterskuddsregisteret 2002
Att.: Arkivverket
Deres ref.: 01/1661 A.48/432 TS
Vår ref.: 2023/3926

Oslo, den 06.11.2023

Avlevering av Etterskuddsregisteret for 2002

Vi viser til tidligere avleveringer fra Etterskuddsregisteret og deres sak 01/1661, som følger vedlagt.

I henhold til bevarings- og kassasjonsbestemmelsene for Skatteetatens elektroniske arkivmateriale pkt. 2 skal det «fra [...] Etterskuddsregisteret bevares samtlige årganger. Disse skal tas ut i form av årgangssnitt av systeme[ts] originale tabeller. Datauttrekk skal fremstilles ett år etter det aktuelle inntektsåret.»

En intern gjennomgang i Skatteetaten har avdekket at det for år 2002 ikke ble produsert avleveringsuttrekk fra Etterskuddsregisteret. Skatteetaten ønsker derfor å overføre gjeldende arkivmateriale og ber i den forbindelse om informasjon knyttet til selve overføringsprosessen.

Vi imøteser deres snarlige tilbakemelding.

Med hilsen

Marcus Liebold

Fagdirektør for dokumentasjonsforvaltning

Skatteetaten

Skattedirektoratet/Virksomhetsstyring/Styring

marcus.liebold@skatteetaten.no
[skatteetaten.no](https://www.skatteetaten.no)

Dataskjema for førsteside

Navn	Beskrivelse
Journaldato	Journalføringsdato
Saksnummer	Saksnummer
Dokumentnummer	Dokumentnummer
Avsendernavn	Avsendernavn
Mottakernavn	Mottakernavn
Saksittel	Saksittel
Dokumenttittel	Dokumenttittel
Dokumentdato	Dokumentdato
Arkivkode	Arkivkode (klassering)
Avskrivingsdato	Avskrivingsdato
Avskrivingsmåte	Enum: TE (Til etterretning), BU (Besvart utgående)

Kjør analyse

Lagre instruksjoner

Utrekk fra dokument

```
1  ▼ i
2  "Journaldato": "06.11.2023",
3  "Saksnummer": "2023/3926",
4  "Dokumentnummer": "undefined",
5  "Avsendernavn": "Marcus Liebold",
6  "Mottakernavn": "Postmottak",
7  "Saksittel": "undefined",
8  "Dokumenttittel": "Avlevering av Etterskuddsregisteret 2002",
9  "Dokumentdato": "06.11.2023",
10 "Arkivkode": "01/1661 A.48/432 TS",
11 "Avskrivingsdato": "undefined",
12 "Avskrivingsmåte": "undefined",
13 "Dokumenttype": "I",
14 "Tilgangskode": "undefined",
15 "Hjemmel": "undefined"
16 }
```

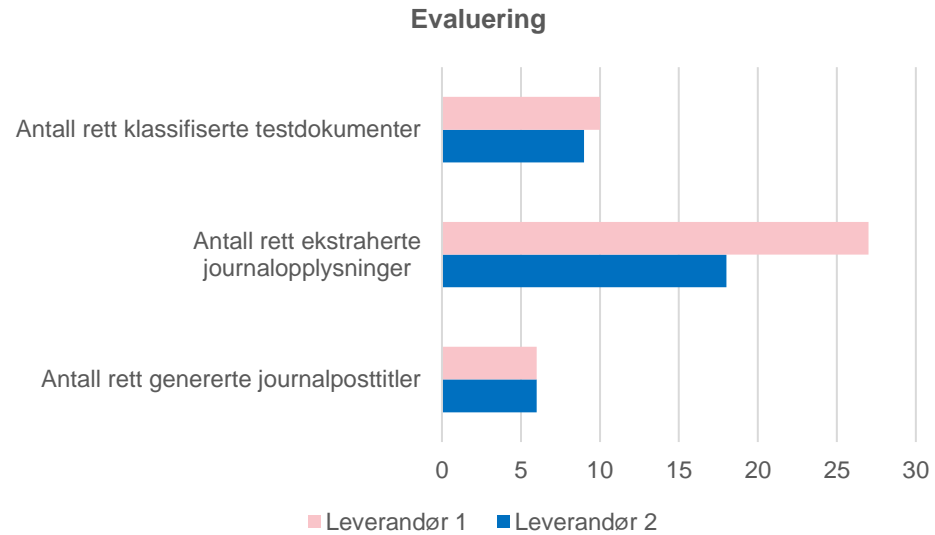
Automatisk identifisering av journalopplysninger:

- Generisk problemstilling som kan anvendes på flere forretningsområder i Skatteetaten
- Saksbehandlingen i Skatteetaten handler i stor grad om å trekke ut data fra dokumenter for enten å be- eller avkrefte påstander
- Journalopplysninger defineres i et dataskjema
- Uttrekk av eksisterende data
- Ikke eksisterende data blir definert som «undefined»
- Kontroll av ekstraherte data kan gjøres ved:
 - å utføre ekstraheringen av dataene to eller flere ganger
 - å be ulike språkmodeller om å ekstrahere dataene og sammenligne resultatene
 - å plukke ut dokumenter som kan kontrolleres av en arkivar

Analyse og generering av forslag til standardiserte journalposttiter:

- Uttrekk av eksisterende data (tittel)
- Identifikasjon av type dokument
- Avsjekk mot Skatteetatens fordelingsliste og skriveregler for enkelthenvendelser
- Forslag til standardisert journalposttittel

<p>Fra: arkivplan23@gmail.com[arkivplan23@gmail.com] Sendt: 31.08.2024 00.06.17 Til: SKD_Innsyn_OEP[SKD.InnsynOEP@skatteetaten.no] Tittel: Innsynskrav</p> <p>Til: Skattedirektoratet Bestillingsdato:31.08.2024</p> <p>Svar på dette innsynskravet ønskes sendt med e-post til: arkivplan23@gmail.com Innsynskravet er bestilt av:</p> <p>Dokumenter: Saksnr: 2022/3938 Dok nr. : 23 Sekvensnr.: 28408/23 Sak: <u>Sivilombudet</u> - Undersøkelse av eget tiltak - om divisjoner og regioner er selvstendige forvaltningsorganer i <u>innsynssaker</u> Dokument: Oversendelse av uttalelse - om Skatteetatens arkivering, journalføring og håndtering av <u>etatsintern</u> korrespondanse Dok.dato: 2023-09-11 Journaldato: 2023-12-02 Saksbehandler: M93878</p>	<p>Innsynsbegjæringer</p>	<p>Tittel: Avhenger hva de ber om innsyn i, legges i saken hvor dokumentene er Tilgangskode: Mappetype: Arkivdel: Klassering: Tilleggsattributter: Elark Base: Merknader:</p>	<p>Dokumentkategori/innhold: 90701B - Krav om innsyn etter Personopplysningsloven, 90801B - Krav om innsyn etter offentlighetsloven, 90903B - Krav om partsinnsyn Merking av innhold: Tilgangskode: Merknader: Se rutine for innsyn</p>
	<p>Innsynsbegjæringer, klager</p> <p>Gjelder ikke klager på innsyn etter offentliglova</p>	<p>Tittel: Klage på avslag om Innsyn <Org.nr/F.nr/D.nr> <NAVN> Tilgangskode: Se rutine Mappetype: Arkivdel: FAG/ADMIN Klassering: M03A - Personopplysningsloven, M03C - Forvaltningsloven, M03D - Folkeregisterloven, M03E - Skatteforvaltningsloven Tilleggsattributter: Ikke aktuelt Elark Base: FAG/ADMIN</p>	<p>Dokumentkategori/innhold: 90904B - Klage på avslag om partsinnsyn 90702B - Klage på avslag om innsyn etter personopplysningsloven Merking av innhold: Tilgangskode: Merknader: Se rutine for Innsyn</p>



Fase 2 – Utvikling av prototype & funn:

- Sammenstilt & syntetisert 10 testdokumenter for klassifisering av arkiv- og journalplikt
- Evaluering av prototypene viser rett klassifisering på 90-100% av dokumentene
- Treffsikkerheten på identifiseringen av obligatoriske journalopplysninger ligger på mellom 47-71%
- Treffsikkerheten på genereringen av standardiserte journalposttitler ligger på 100%



Tilgjengeliggjøring / implementering:

- Utvikling av en agent i Microsoft Copilot Studio
- Tilgjengeliggjøring for Skatteetatens ansatte som p.t. har en lisens for Microsoft 365 Copilot (220-250 stk.)
- Agenten skal være tilgjengelig på alle arbeidsflater som Microsoft 365 Copilot er tilgjengelig på (eksempelvis Teams, Word, Outlook m.fl.)
- Agenten skal operere i brukerens kontekst og hensynta brukerens tilganger
- Utrulling av arkiveringsassistenten for alle ansatte i Skatteetaten ila 2. halvår 2025

Lagt inn i Alle sammen

 IT, redaksjon
21. mai • Redigert

delt 1 gang • Sett av 2 089

Skatteetaten inviterer næringslivet til å utvikle en arkiveringsassistent basert på kunstig intelligens

Arkivering og journalføring kan være utfordrende. Derfor fikk vi bl.a. negativ omtale rundt vår fagvurderinger *Skattemyndighetene* bla (bt.no)

Etter at det ikke opplever krever t

Kan en Fagdirektiv initiativ modern



– Vi er veldig fornøyd med responsen i den første fasen av anskaffelsen. Det var 12 forslag som kom inn, og alle var gode, sier Marcus Liebold Skatteetaten. Foto: Skatteetaten

Første i sitt slag

Konkurransen er den første i sitt slag fra Skatteetaten.

I stedet for å inngå avtale med én leverandør, som skal levere en fullverdig løsning basert på kravene, gjør etaten nå en såkalt førkommersiell anskaffelse. Metoden er utviklet for å implementere forsknings- og utviklingstjenester, og skal passe godt om det er behov for å utvikle løsninger som ikke finnes i markedet fra før.

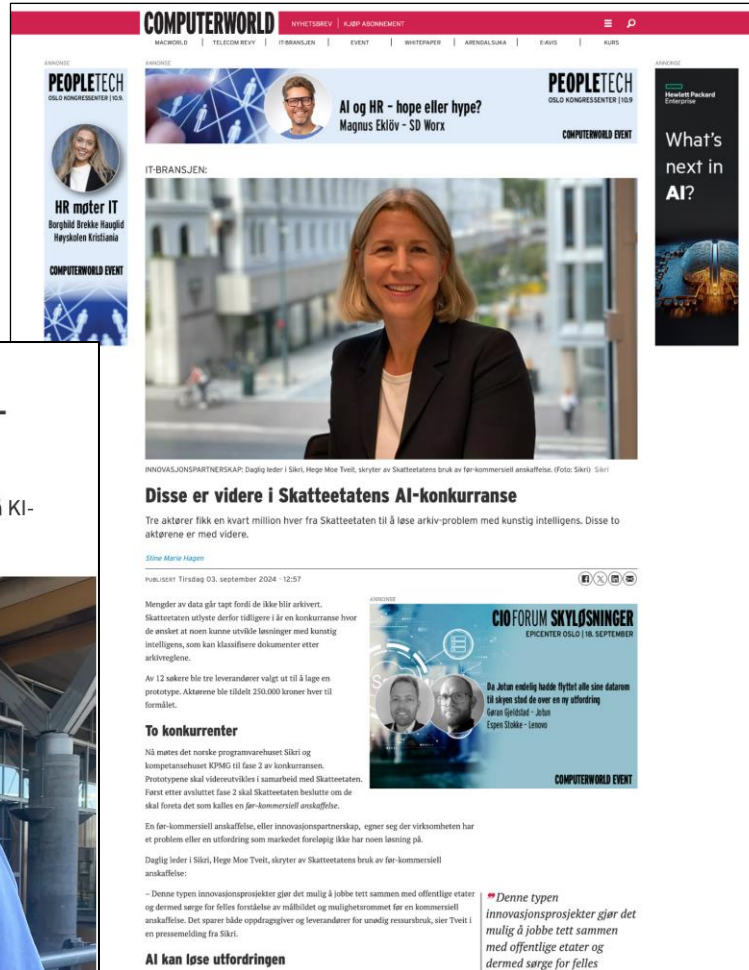
DIGITALISERING OG OFFENTLIG IT

Disse får penger for å løse arkiv-floke

Tre selskaper får en kvart million hver for å utvikle en prototyp på KI-støttet arkivering: KPMG, Sikri og Best of breed.



er usikkert. – For å være helt ærlig, så er dette så viktig for Best of breed h, daglig leder i Best of breed Foto: Privat



The screenshot shows a Computerworld article titled "Disse er videre i Skatteetatens AI-konkurranse". The article discusses a competition where three companies (KPMG, Sikri, and Best of breed) have been selected to develop a prototype for AI-supported archiving. The article includes a photo of Hege Moe Tveit, the daily leader at Sikri, and a quote from her. The article also mentions that the competition was part of a pilot project for a commercial procurement process. The article is dated Tuesday, September 3, 2024, at 12:57.

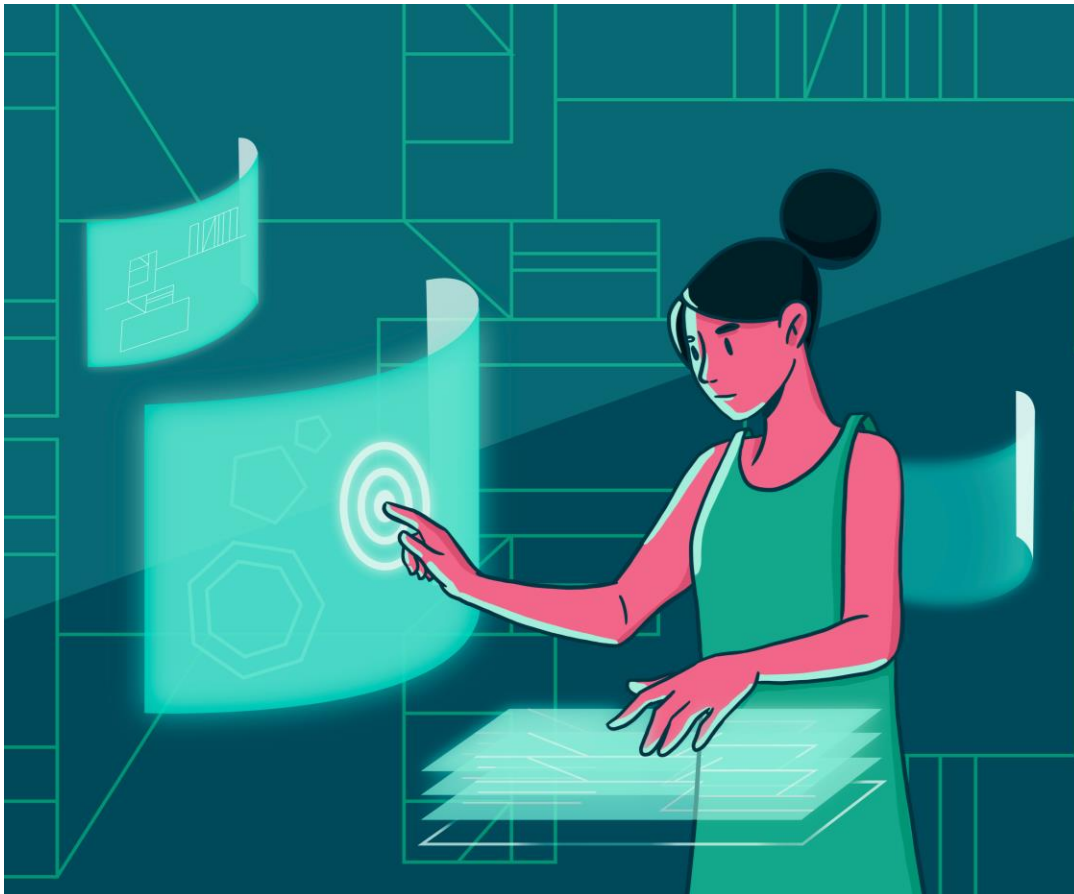
Kommunikasjonsstrategi for eksperimentet:

- Behov for å nå ut med informasjon om eksperimentet:
 - internt i Skatteetaten (intranett, Viva Engage mm.)
 - eksternt ([nettsteder](#), Arendalsuka m.fl.)

«En utrolig spennende case som er gøy å jobbe med!»

«Vi kommer til å satse på utviklingen av et slikt verktøy uansett om vi vinner konkurransen eller ei.»

«Dette er måten vi egentlig ønsker å jobbe på med utvikling av programvare.»



Hva betyr tiltaket for Skatteetaten?

- Effektivisering av arbeidsoppgaver innen dokumentasjonsforvaltningen med utvidelse på flere områder
 - Innsynshåndtering
 - Tilgangsvurdering
 - Postfordeling
- Frigjørelse av kapasitet til kjerneoppgaver
- Etterlevelse av lovpålagte krav om arkivering og journalføring av dokumentasjon
- Omdømmebygging



Skatteetaten



Marcus Liebold

Fagdirektør for dokumentasjonsforvaltning
Skattedirektoratet – Styring og utvikling
marcus.liebold@skatteetaten.no