

# Proof of Concept:



POC av egnethet for bruk av Anzyz kunstig  
intelligens løsning på Kristiansand kommunes data

Juni 2019

# Innhold



- Bakgrunn
- Prosjektbeskrivelse
- Bærekraft
- Teknologien
- Casene
- Resultat
- Veien videre
- Hovedprosjekt
- Konklusjon
- Kontakt

## Bakgrunn:

Kristiansand kommune har gjennomgått både kvantitativ og kvalitativ analyse av data i de ulike journalsystemer kommunen har. Dette arbeidet er publisert i rapporten *Sosialhjelpsanalysen*.

«Tanken var at vi kunne benytte rapporten og grunnlagsmateriale til denne som en slags «fasit» og teste om deres løsning finner det samme når den behandler journalopplysninger – både kvantitative felt og tekstfeltene i journalsystemet» uttrykker Kristiansand kommune i mail av 15.08.18.

I tillegg til disse, ønsket Kristiansand kommune å undersøke mulighetene innenfor *helsedata* i forbindelse med pasientsikkerhetsprogrammet: *Forebygging av fall i helseinstitusjoner og forbedringsarbeid*.

I samarbeid med Kristiansand kommune, ble det definert ulike caser for å teste egnetheten for bruk av Anzyz kunstig intelligens løsning på kommunens data, i form av en *Proof of Concept* (POC). Datagrunnlaget ble basert på journalopplysninger fra *Profil* og *Acos*.

# Prosjektbeskrivelse

POC-en ble delfinansiert av Innovasjon Norge. I søknaden om prosjektmidler fremgår følgende begrunnelser for POC-en:

- Å teste ut en innovativ anvendelse av kunstig intelligens innenfor et nytt marked, noe som vil kunne gi Anzyz erfaringer som kan lede til ny anvendelse, nye markeder og kunder.
- Beslutningsstøtte og høyere kvalitet i beslutninger både rundt enkelttjenester og på overordnet nivå for Kristiansand kommune. Bedre kvaliteten på data og tilgjengeliggjøring av informasjon på en ny måte, som vil gi bedre kvalitet på tjenester til innbyggere og brukere.
- Effektiv ressursutnyttelse gjennom bedre beslutninger, muliggjøre digitalisering av arbeidsprosesser og målrettede tjenester innenfor kommunens tjenestetilbud.
- Kompetanse og erfaring med bruk av kunstig intelligens innenfor kommunesektoren - også krav knyttet til GDPR.

**Målet for POC-en** er å avgjøre om datamaterialet kan besvare casene fra Kristiansand kommune. Man ønsker å avgjøre om et hovedprosjekt er gjennomførbart, samt avklare hvordan et slikt prosjekt kan gjennomføres. Resultatene av POC-en beskrives i en rapport.

## Bærekraft:

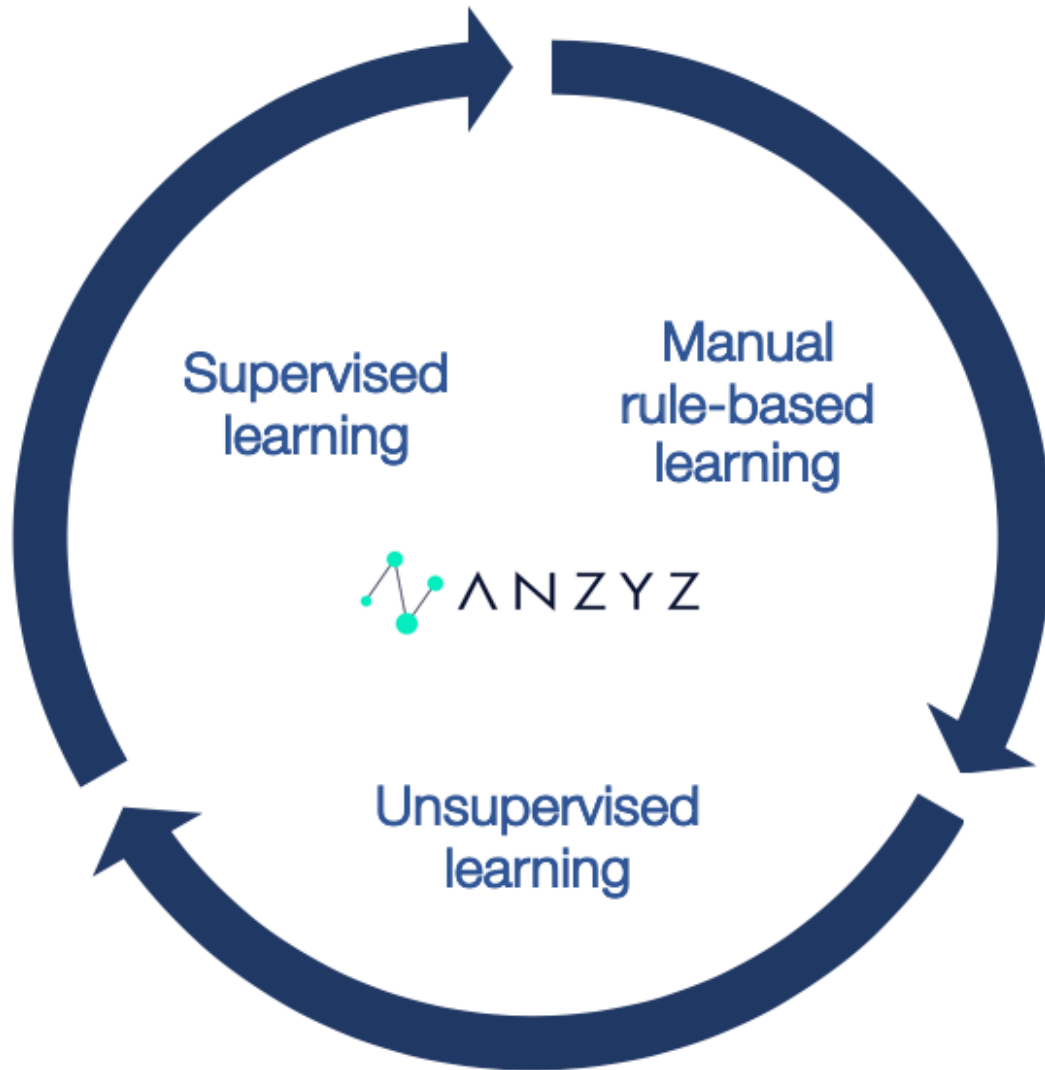
- Etikk i bruk av kunstig intelligens er viktig. Anzyz er medlem av European Artificial Intelligence Alliance som er etablert av den Europeiske Kommisjon som et diskusjonsforum for utviklingen, innvirkningen og betydningen av kunstig intelligens i samfunnet.
- Ett av målene med POC-en for Kristiansand kommune er å kunne vise hvordan kunstig intelligens har en samfunnsnyttig effekt gjennom blant annet effektivisering/automatisering av manuelle prosesser, noe som vil redusere tidsbruk og dermed øke kvaliteten på tilbudet til kommunens borgere.

# Teknologien



Anzyz trekker kunnskap om setninger, fraser og ord direkte ut fra ustrukturerte data gjennom en selvlærende mekanisme basert på korpuslingvistikk\*. Dermed fanges språket dynamisk, og ikke basert på forhåndsdefinerte nøkkelordlister eller språkmodeller. Teknologien kombinerer fordelene ved nøkkelordbaserte, regelbaserte og maskinlæringsbaserte systemer, og eliminerer dermed sentrale svakheter ved manuelle analyser.

\* «Korpuslingvistikk er en språkvitenskapelig metode som baserer seg på automatiserte analyser av store mengder språkdata i elektronisk form (korpus). Korpuslingvistikken brukes mye ved språkprosessering og tekstanalyse og som forskningsmetodikk innenfor samfunnsvitenskap og humaniora» fra <http://uni.no/nb/topics/corpus-linguistics0/>



- Basert på innovasjon og forskning
- Integrerer 3 forskjellige fremgangsmåter
- Forstår kontekst – neural probability networks
- Forstår ordenes *funksjon* og *rolle*
- Data + algoritme = AI base
- AI-base: ustrukturert data er blitt strukturert

# Casene



I samarbeid med Kristiansand kommune, ble det definert ulike caser for å teste egnetheten for bruk av Anzyz kunstig intelligens løsning på kommunens data, i form av en *Proof of Concept* (POC). Datagrunnlaget ble basert på journalopplysninger fra *Profil* og *Acos*.



# Caser, basert på *Sosialdata* fra Acos:

1. Acos har en systemfeil som teller feil i antall barn som finnes i en husholdning. Det er ønskelig at Anzyz kan gi en korrekt rapport over følgende:
  - a. Hvor mange husholdninger som har barn blant sosialhjelpsmottakerne i 2017.
  - b. Hvor mange barn totalt som lever i husholdninger som mottok sosialhjelp i 2017.
2. Av alle mottakerne i 2017 var det en andel som stod oppført med norsk statsborgerskap parallelt med at de også stod oppført med "født i Norge". En ikke uvesentlig andel av disse stod feiloppført, og var faktisk født i et annet land.
  - a. Kristiansand kommune ønsker at Anzyz skal, basert på skriftlig historikk i saken, identifisere brukerne som stod feiloppført.
3. En viss andel av ungdommen under 30 år var uten fast bopel (UFB) i 2017.
  - a. Kristiansand kommune ønsker at Anzyz skal identifisere disse og gi en prosentandel av gruppen totalt sett som i hele eller deler av 2017 var UFB.

# Caser, basert på *Helsedata* fra Profil:

1. Kristiansand kommune ønsker å hente ut informasjon i brukerjournaler knyttet til *fall*. Dette er knyttet til forbedringsprosjektet *Forebygging av fall i helseinstitusjoner*.
  - a. Hvor mange som falt
  - b. Fall med og uten skade
  - c. Fordeling av fall mellom avdelinger
  - d. Hvor mange dager det går mellom fall

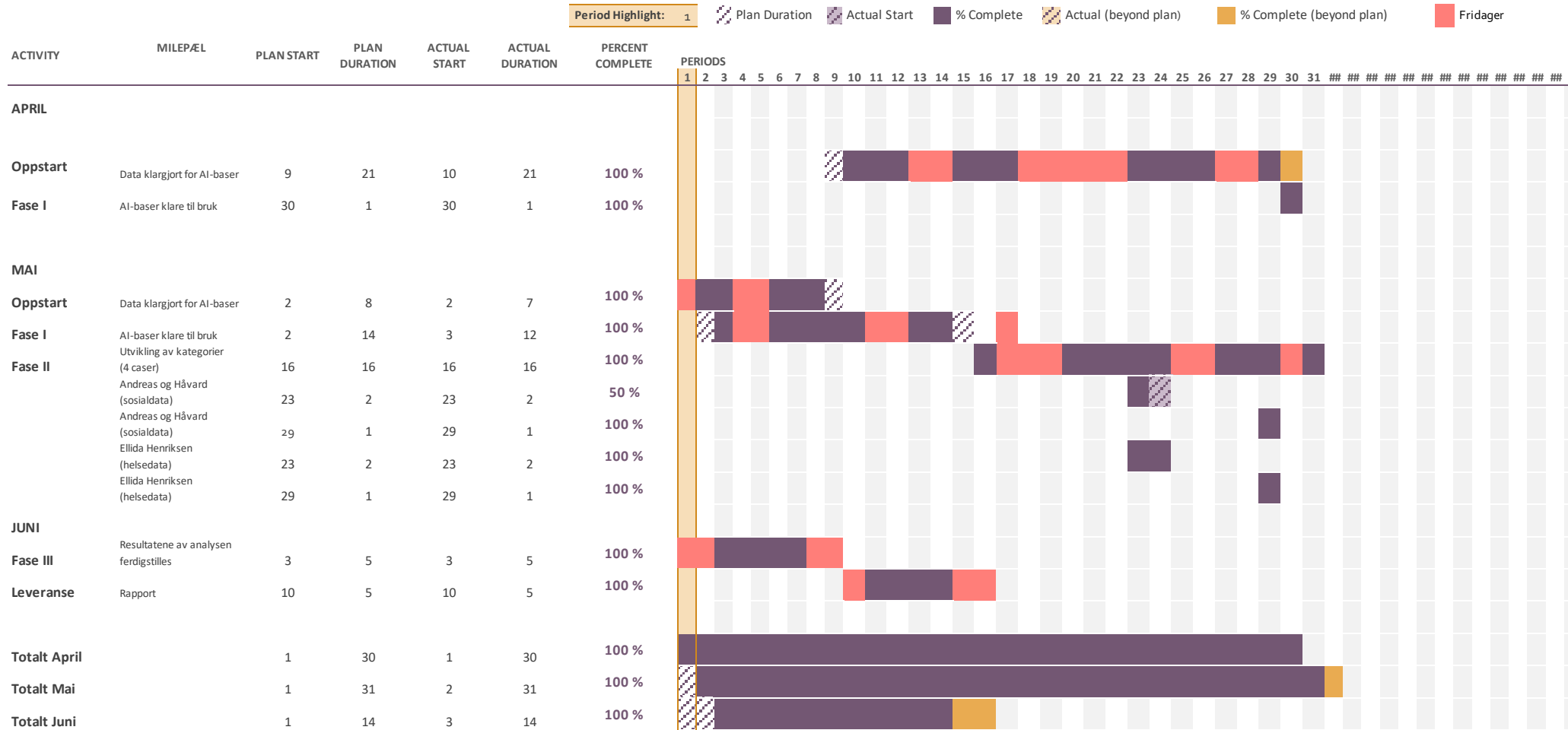
# Resultat



Følgende kapittel ser nærmere på resultatene for casene, med kommentarer til de ulike funnene, samt metodikk.

# Gjennomført prosjektplan:

## POC Kristiansand kommune



## 2. Kristiansand kommune ønsker at Anzyz skal, basert på skriftlig historikk i saken, identifisere brukerne som stod feiloppført:

Av alle mottakerne i 2017 var det en andel som stod oppført med norsk statsborgerskap parallelt med at de også stod oppført med "født i Norge". En ikke uvesentlig andel av disse stod feiloppført, og var faktisk født i et annet land. Basert på datagrunnlaget i journalene, ble flere identifisert som feiloppført basert på søk i dokumenter tagget med *Norsk opprinnelsesland*.

### Se eksempler:

- Kom til Norge i februar 2017.
- Musikk lærer i hjemlandet.
- Født i Irak, bodde mange år i Iran.
- Hun kom til Norge i 2014
- De kom til Norge i 2013.
- NAVN sier han kom til Norge i september og hadde med seg oppsparte midler.
- NAVN kom til Norge i 2003.
- NAVN kom i 2001.

### 3. Kristiansand kommune ønsker at Anzyz skal identifisere disse og gi en prosentandel av gruppen totalt sett som i hele eller deler av 2017 var UFB:

Basert på datagrunnlaget i journalene, ble flere identifisert som registrert med fast bopel, og var dermed feiloppførte.

#### Se eksempler:

- Står for tiden uten bolig.
- Søker er nå UFB.
- Bruker står uten bolig nå etter et samboerforhold som er avsluttet.
- Hun forteller at hun er uten bolig nå, har vært det i over en og en halv måned mens hun har leitt etter egen bolig.
- Er nå uten bolig, flyttet ut i begynnelsen av desember, men betalt ut januar. Leter ny bolig.
- Han har stått uten fast bolig i 1.5 måned, men har fått bo her og der hos venner.
- Personal har senere fått vite at han oppholder seg i GATENAVN og fortsatt er ufb.

for tiden ufb  
uten fast bopel  
er ufb  
fortsatt uten bolig  
uten fast bosted  
ufb  
ikke har bolig  
uten fast bolig ,  
nå uten bolig ,  
uten bolig ,  
ufb  
ingen sted å bo  
uten bolig  
uten fast bolig  
nå ufb

## Metodikk for søk i Profil:

### FAKTA OM AI-BASEN:

Totalt antall logger: 6 293 583

Antall unike fraser: 619 221

Uttrekk: logger fra 01.01.2010 og frem til i dag.

- Hentet inn alle logger fra januar, 2010 til mai, 2019.
- Kjørte så maskinlæring på all data og satt opp systemet.
- Etablerte kategorier på «Fall» og «Skade som er relatert til fall».
- Gikk deretter gjennom disse kategoriene sammen med domeneekspert, Ellida Henriksen, for å kvalitetssikre data.
- Utfordringen med Profil har hovedsakelig vært å få tak i all den aktuelle metadata til loggene.

# Definisjon av fall

I samråd med domeneekspert, Ellida Henriksen, ble *fall* definert etter følgende parameter:





## Kommentar:

- Potensiell feilkilde: at samme fall muligens kan telles flere ganger, derav benevnelsen *omtalte fall*.
- For å motvirke at fall ble dobbeltregistrert, ble det lagt til mot-eksempler på fraser. Eksempler:  
har tidligere falt / falt forrige ... / har falt før / osv.
- Dette konseptet kan videreutvikles i et hovedprosjekt, slik at man oppnår enda høyere treffsikkerhet.

# b) Omtalte fall med og uten skade

Skade ble definert i samråd med domeneekspert, Ellida Henriksen:

Skade i forhold til fall



Showing 250/250 words in result

# c) Fordeling av omtalte fall mellom planområder:

2019

Anzyz gjorde følgende funn i analysen av fordeling av omtalte fall mellom planområder:

2019

Flest fall prosentvis:	Antall fall:	Totalt antall logger:	Prosent:
Helseveiledende samtale		52	
Overflytningsnotat		1937	
Sengesensor		6	

Flest fall i antall:

Helse	214960
Legekonsultasjon	8720
Bevege seg inne og ute	14487

## Kommentar:

- Innenfor fagsystemet, viser *område* til hvilket planområde det er, slik som for eksempel:

Motivasjonsfase / Treningstilbud- hjerneslag / Helseveiledende samtale / Sjekklistor

- Fagsystemet viser altså ikke nødvendigvis til konkrete avdelinger. Det ble forsøkt å kryssjekke ansatte fra loggsystemet opp mot register i Økonomisystemet for å definere konkret avdeling. Dette viste seg å være vanskelig og tidkrevende, ettersom det fantes svært mange ulike muligheter.
- I tillegg, hadde samme område flere forskjellige benevnelser. Her er noen få eksempler av variasjoner av Presteheia omsorgssenter:

Presteheia avd. 2 / Presteheia avd. 2 a sterk / Presteheia natt / Presteheia bygg/renhold  
/ Presteheia bolig / Presteheia avd C / Prestheia 83

### Forslag til løsning:

- Forbedring av datagrunnlaget ved en manuell gjennomgang av alle steder man er ansatt og koble det til et omsorgssenter.
- Koble alle variantene av Presteheia omsorgssenter, slik at man kan se på forskjeller mellom omsorgssentrene. Eventuelt kan man ha en gjennomgang av all data som er lagret i databasen og finne en alternativ måte å mappe denne dataen på.

## d) Hvor mange dager det går mellom omtalte fall

Kommentar:

- Funnene viser at datagrunnlaget i liten grad egner seg til å svare på dette spørsmålet. Ettersom tidspunktet for loggføringen av *fall* ofte ikke svarer til når fallet faktisk oppstod. Eksempler kan være at *beboer falt i går* eller *fall for to år siden*, osv.
- Disse utfordringene kan eventuelt forbedres i et hovedprosjekt.

# Hovedprosjektet

Det har tidligere vært diskutert mulighetene for å utforske følgende caser for et hovedprosjekt:

- Bedre styring og ledelse med bedre tilgjengelig styringsdata med god kvalitet.
- Digitalisere manuelle arbeidsprosesser, eksempelvis kombinere AI med robotteknologi.
- Finne de pasient- og brukerhistoriene som vi lykkes med og finne frem til risikogrupper/målgrupper som en skal innrette seg mot.

# Konklusjon



Under arbeidet med POC-en, ble det klart at domeneeksperter bør være med under datauttrekk (ref. Acos).

Det vil i fremtiden være viktig for Kristiansand kommune å ha tilgang på egne data. Som læring fra erfaringene gjort med Acos, vil det være vesentlig at kommunen i fremtiden har teknisk enkel og rask tilgang til sine data.

Det anbefales å involvere Kristiansand kommune i dialogen med Acos for å få tilgang til det resterende Sosialdatagrunnlaget.

Funnene i rapporten viser at det er et stort videre potensial for bruk av kunstig intelligens innenfor Profil og helsedata.