

**Til:** Digitaliseringsminister Nikolai Astrup

**Fra:** Norconsult Informasjonssystemer, ved adm. dir. Ola Greiff Johnsen

**Dato:** 8. oktober 2019

---

## **Digitalisering av offentlig sektor: Innspill forbedringsområder**

Vi viser til møte med digitaliseringsminister Nikolai Astrup og Bærum Høyre hos Norconsult i Sandvika, 4. sept. 2019. I møtet ble vi oppfordret til å komme med innspill til sentrale forbedringsområder knyttet til digitalisering av offentlig sektor.

Under har vi oppsummert de mest sentrale forbedringsområdene, slik vi opplever det ut fra et næringslivsperspektiv.

- 1) Behov for å løfte datakvalitet og arbeide med bedre struktur i basisregisteret Matrikkelen.
- 2) Potensial for å hente større gevinster fra maskinlæring og kunstig intelligens.
- 3) Potensial for effektivitetsgevinster ved bedre tilgang og deling av data på tvers av virksomheter i offentlig sektor.

Dette er problemstillinger som vi har beskrevet mer detaljert i dette notatet. Ta gjerne kontakt med undertegnede dersom dere ønsker ytterligere informasjon, eller om det er annet vi kan bidra med.

Med vennlig hilsen  
**Norconsult Informasjonssystemer AS**

Ola Greiff Johnsen

Administrerende direktør

## 1 Matrikkelen

Matrikkelen er sammen med Folkeregisteret og Enhetsregisteret definert som nasjonalt basisregister. Matrikkelen er dermed et sentralt register som benyttes i mange fagsystem. Ved kommunereformen ser man at det er mange hundre fagsystem som berøres, alt fra NAV, helse, SSB, posten, kommunale arkivsystem, vegdatabanken, etc. Dessverre er informasjonen i Matrikkelen av en helt annen kvalitet enn informasjonen i de to andre nasjonale basisregistrene.

Fra Norconsults perspektiv jobber vi med fagsystem for selve matrikkelføringen samt systemene for kommunal forvaltning som baserer seg på hvem som til enhver tid eier en eiendom.

Matrikkelen har utfordringer både mhp struktur og datakvalitet. Fra vårt synspunkt er følgende problemstillinger sentrale:

**Struktur og arkitektur:** Matrikkelen håndterer verden i to dimensjoner. I den virkelige verden foregår det eiendomstransaksjoner i tre dimensjoner, for eksempel ved at en utbygger eier grunnen under bakken og bygger et parkeringshus, mens en annen aktør eier kjøpesenteret som bygges over parkeringshuset.

**Historikk:** Matrikkelen inneholder rimeligvis historikk, men dette tilbys ikke som tilgjengelig informasjon for fagsystemene. De fleste løsninger som bygger på matrikkelen har behov for historikk knyttet til en eiendom. Dette gjør at hver leverandør har laget sin egen løsning for håndtering av historikken. Dette har åpenbare tekniske svakheter og det ville vært en bedre løsning om matrikkelen hadde et grensesnitt for dette.

**Kvalitet i matrikkelen.** Kvaliteten på dataene i matrikkelen er en stor utfordring og begrenser mulighetene for digitalisering i offentlig sektor. Når man er skeptisk til kvaliteten i basis-registeret vil dette være til hinder for å innføre automatiske prosesser for kommunal saksbehandling.

I en situasjon der sentrale tjenester som brann og ambulanse sentraliseres er det avgjørende at grunnlagsdataene har god kvalitet. I dag er det fortsatt mange kommuner som ikke har vegadresser, men benytter matrikkeladresser selv om dette har vært et krav siden medio 2018

Noen eksempler på kvalitet i matrikkelen:

- Nibio og SSB har gjennom tidligere undersøkelser vist at ti prosent av landets store utmarkseiendommer mangler eiendomsgrenser og eier i Matrikkelen. Til sammen utgjør disse eiendommene rundt en femdel av Norges landareal.
- Tall fra Kartverket viser at fundamentale opplysninger som gårds- og bruksnummer mangler ca 3,5% matrikkelenhetene i Matrikkelen (I følge tall fra Kartverket fra 2017 er 88.208 matrikkelenheter registrert uten teig)
- Et stort antall eiendommer mangler eiendomsgrenser. (I følge tall fra Kartverket høsten 2017 hadde 127.131 matrikkelenheter fiktive grenser, dette er ca 5% av eiendomsmassen)
- Ved forrige runde av kommunesammenslåinger skulle to sammenslåtte kommuner ha felles gebyr-regulativ. Det ble bestemt at areal på bygningsmassen registrert i matrikkelen skulle danne grunnlaget for vannmålergebyret. Da dette skulle settes i drift oppdaget man at alle bygg i den ene kommunen var registrert med areal 100 m2. Når så fundamentale feil oppdages vil tiltroen til kvalitet på andre deler av informasjonen i Matrikkelen være liten.
- Ved utarbeidelse av byggesøknader som gjelder endringer eller ombygging av eksisterende bygg, så er ofte ikke arealopplysningene i matrikkelen av en slik kvalitet at de kan benyttes som direkte grunnlag i søknaden. Dette fører til merarbeid ved at man må gå inn i flere saker i byggesaksarkivene og det fører ofte til ekstra dispensasjonssøknader. Dette øker kostnader for tiltakshaver og det gir lengre og mer omfattende saksbehandling i kommunene.

Kartverket har igangsatt et prosjekt som heter «Masterplan matrikkelen» som skal se på disse problemene. Vi håper at også departementet engasjerer seg i dette og bidrar til at prosjektet får tilført nok kompetanse og midler slik at dette sentrale registeret oppnår en tilfredsstillende kvalitet.

## 2 Maskinlæring og kunstig intelligens

Norconsult Informasjonssystemer har erfaring fra AI og maskinlæringsprosjekter både i privat og offentlig sektor. Vår erfaring er at sammenfallet av store mengder data, billig prosesseringskraft og modne rammeverk, gjør at mange virksomheter og sektorer kan hente ut potensielt store verdier ved bruk av denne type teknologier.

Markedet er nå i en fase hvor mange aktører har vært gjennom de første utprøvingene. Flere har gjennomført konseptutprøvinger - mange med positive resultater - og et mindre antall har satt i produksjon de første maskinlærte modellene. Våre råd om hvordan det kan legges til rette for at offentlig sektor kan hente ut mer av mulighetene som ligger i AI og maskinlæring bygger på erfaringer fra alle disse fasene hos både store og små dataeiere.

### 2.1 Tilgang til data

Alle som jobber med maskinlæring vil være enige i mulighetene for å lykkes med AI/ML står og faller på tilgangen på nok data av god kvalitet. Status slik vi ser det, er at offentlige data ofte er låst inne i siloer, at mange offentlige aktører har tilgang på for lite data og at den juridiske situasjonen rundt bruk av dataene er uavklart.

Gjennom erfaringer fra flere store organisasjoner, har vi innsett verdien av å investere i å gjøre all data tilgjengelig i en datasjø, slik at data kan sammenkobles og utnyttes på tvers av avdelinger i organisasjonen. Dette er et krevende steg, men er avgjørende for å gjøre det lettere å utforske mulighetene som finnes i dataene, og senere å høste verdier gjennom maskinlæring og AI. Hvis alle konseptutprøvinger og alle prosjekter må gjøre jobben med å hente inn data selv, blir terskelen for stor og innovasjonen hindres.

De offentlige aktørene som sitter på store, strukturerte datasett, slik som for eksempel NAV og Skatteetaten, har allerede kommet langt i etablering av datalager med tilhørende datasjø og har dedikerte analyseavdelinger. De mindre etatene har ikke de samme mulighetene. For offentlig sektor vil mye data også være nyttig på tvers av virksomheter, og vi anbefaler at det legges vekt på å fortsette arbeidet med å utrede hvordan en kan etablere en datasjø for deling av data på tvers av virksomheter i offentlig sektor. Vi er klar over alle de organisatoriske, tekniske og juridiske vanskelighetene som gjør at mange tar til orde for mer løst koblede alternativer, men vi er altså av den oppfatningen at disse alternativene ikke er fullgode.

Mange offentlige data ligger låst inne i systemer som gjør at det er svært omfattende oppgave å hente dem ut på en måte som gjør dem egnet til maskinlæring. Det kan være at systemene er programvare hvor leverandøren ikke tilbyr grensesnitt for å hente ut data, eller det kan være at dataene ligger i generiske datamodeller som gjør at det ikke er lett å hente dem ut "fra baksiden". Det ligger også problemstillinger knyttet til sikkerhet som gjør at veien til maskinlæring er kronglete. For et enkeltstående prosjekt kan dette være hinderet som gjør at man ikke kommer videre. Dette problemet kan håndteres ved å utarbeide veiledninger som offentlige virksomheter kan benytte ved anskaffelser av IT-system. Virksomhetene bør her pålegges å stille spesifikke krav til leverandører om å levere API som gjør at data fra disse systemene kan utnyttes i større grad.

Mange data er utilgjengelige for maskinlæring grunnet juridiske problemstillinger. Den mest åpenbare utfordringen er avveiningen mot personvern hensyn og GDPR. Et grunnleggende problem med bruk av persondata til maskinlæring er at dataene ikke er hentet inn på en måte hvor innbyggeren har gitt samtykke til bruk av data til dette formålet. Dette er problemstillinger som må adresseres om man ønsker full nytte av denne teknologiens muligheter i offentlig sektor.

Noen mener at dette kan løses gjennom anonymisering, obfuskering (program for å skjule data) og aggregering, men vår vurdering er at slike mekanismer ikke vil fungere i det lange løp. Verdien av maskinlæring ligger i å utnytte

svake mønstre i data. Hvis man obfuskerer for sterkt, mister man disse mønstrene. Hvis man gjør det for svakt, kan sikkerheten brytes.

Andre juridiske områder kan være sikkerhetslovgivningen, lov om helseregistre, hensynet til forretningssensitive data. Det er åpenbare grunner til at man har lovgivning som beskytter data, men dette er premisser man må være klar over før man starter maskinlæringsprosjekter.

På dette området foreslår vi:

- Juridiske avklaringer knyttet til bruk av data ved at Digitaliseringsdirektoratet presser på for at det etableres normer for hvordan data kan brukes til maskinlæring. I noen tilfeller gjennom lov- og forskriftsendringer.
- Samtykke til AI/ML fra innbyggere gjennom en sentral ordning eller ved at de enkelte dataeiere i offentlig sektor oppdaterer sine samtykkeordninger.
- Krav til å gjøre data enkelt tilgjengelig fra systemer i alle nye IT-prosjekter og etablering av standardiserte API-er for datatilgang i eksisterende systemer.
- Fortsette arbeidet med datadeling i offentlig sektor, og prioritere å gjøre data tilgjengelige i en datasjø fremfor mer desentrale løsninger.

## 2.2 Gjennomføring av ML/AI-prosjekter

Nest etter tiltak knyttet til tilgang til data, mener vi at man må se på hvordan man jobber med innovasjon, utprøving, utvikling og implementering av AI/ML. Teknologien har noen særtrekk som krever egne tilnærminger: lav inngangsterskel, avhengighet til data, store behov for tverrfaglig samarbeid og nye IKT-krav.

Summen av disse særtrekkene fører til et fenomen hvor man får lovende resultater fra konseptutprøvinger, men ender opp med løsninger som blir svært krevende å ta videre fra konseptfase til produksjon. I konseptfasen har man sett eksempler på at man tar for lett på etiske problemstillinger, juridiske hindringer, konsekvensen av at AI/ML introduserer nye teknologier i produksjonsmiljøene, behovet for retrening av modeller, og behov for kompetanse knyttet til dette. I tidligfase må man derfor være både strategisk og taktisk når man starter og følger opp prosjekter. Vår anbefaling er å tenke langsiktig og fokusere på anvendelser som har stort verdipotensial. Lovende konsepter med marginal nytte/kost kan ellers ende opp med å koste mer enn verdien de er ment å realisere.

Siden AI/ML er sterkt avhengig av data, er det ofte behov for å justere retningen på prosjektene etterhvert som man avdekker den reelle tilstanden i tilgjengelige data. Kanskje er dataene av for dårlig kvalitet for å realisere den opprinnelige ideen, men godt egnet til en lignende ide. På grunn av dette, må man i høy grad ha en direkte dialog mellom fagfolkene på AI/ML (data scientists), fagfolk fra virksomheten og ledelse/sponsorere/eiere. Dette tverrfaglige samarbeidet betyr at sentrale ressurser må sette av betydelig tid. Flere av de offentlige virksomhetene vi har levert oppdrag til har innsett behovet for egne tilnærminger for å lykkes med AI/ML, og har tilpasset sine prosesser og rammer. Disse erfaringene må deles slik at offentlig sektor ikke kaster bort ressurser på å starte små, dårlig forankrede og halvveis prioriterte prosjekter som mislykkes etter konseptutprøvningsfasen.

På dette området foreslår vi:

- Kompetanseheving hos virksomhetsledere, prosjekteiere og gevinsteiere i relevante gjennomføringsmetoder og etablering av prinsipper for evaluering og styring av AI/ML-prosjekter.
- Etablering av et felles metodeverk eller som en spesialisering av prosjektveiviseren fra Difi.
- Systematisk publisering av erfaringer fra analyse, kartlegging, konseptutprøvinger, utvikling og implementering av AI/ML. Forslag til temaer kan være: business case, juridiske vurderinger, valgte AI/ML-metoder, datagrunnlag.
- Utarbeidelse av retningslinjer for etiske vurderinger på områder som er felles for mange offentlige aktører.

## 2.3 Sourcingstrategier

I motsetning til hva mange tror, stiller maskinlæringsmodeller store krav til drift og vedlikehold. En oppfatning er at siden algoritmene lærer av data, så er det ikke behov for å oppdatere systemene for å lære å håndtere nye tilfeller. Det motsatte er tilfellet: maskinlærte løsninger er datadrevne, og endringer i dataene har mye større konsekvenser

enn i tradisjonelle systemer. På en side er det nettopp denne dynamikken man er ute etter for å ta ut verdipotensialet i mønstre som man ellers ikke kunne fanget opp. Man må samtidig overvåke løsningene for å sikre at modellene ikke påvirkes på uventede, negative måter av endringer i data. Det kan være naturlige endringer i datagrunnlaget, eller det kan være som følge av ondsinnet manipulasjon. Det kan også være tilpasninger i brukermassen som følge av innføringen av AI. Vi skal ikke gå gjennom alle aspekter av dette komplekse området, men konsekvensen er at drift av maskinlærte løsninger krever kontinuerlig tilgang på kompetanse. Når man kjøper inn slike løsninger, må man derfor i mye større grad enn tidligere sikre tilgang på spesialister i forvaltnings- og driftsfasen. Enten må man bygge opp kompetansen selv, eller knytte den til seg gjennom forpliktende ressursavtaler.

Et annet alternativ er å overlate større virksomhetsprosesser til private aktører og la disse løses utfordringene med hvordan de optimaliseres gjennom innovativ bruk av AI/ML. Dette kan for eksempel gjøres ved at man kjøper tjenester gjennom BaaS-avtaler (Business as a Service) eller ved utdeling av konsesjoner. På denne måten slipper det offentlige selv å bygge opp kompetanse på organisering, styring og gjennomføring på et område som er i svært rask utvikling og overlater til markedet å velge de konkrete løsningene og sørge for å være leveransedyktig.

#### På dette området foreslår vi:

- Det offentlige må bygge opp kompetanse på AI/ML i sine utviklings, forvaltnings- og driftsorganisasjoner. Minimum er å bygge tilstrekkelig bestillerkompetanse.
- Drift- og forvaltningsavtaler for AI/ML-løsninger må være langsiktige og ha forpliktende ressursavtaler.
- Krav om at AI/ML-kode leveres som open source. Dette med tanke på erfaringsdeling, standardisering, kompetanseoverføring, gjenbruk og transparent.
- Vurder å sette ut hele virksomhetsområder til privat sektor gjennom BaaS-avtaler eller konsesjonsordninger.

### **3 Digital verdikjede**

Digitalisering handler om å digitalisere hele verdikjeden. Det er lite poeng i at en leverandør er best på digitalisering om all informasjon inn og ut er analog. I vårt prosjekt E39 Kristiansand Øst- Mandal Vest er det krav fra Nye veier om heldigitale leveranser. Entreprenøren AF, byggherren Nye Veier og Norconsult som rådgiver jobber sammen for å digitalisere hele kjeden. I E39-prosjektet er det konkrete eksempler på at det offentlige krever leveranser på papir, noe som skaper brudd i den digitale verdikjeden. Vi ønsker at departementet fortsetter å ha fokus på offentlig sektors bidrag til digitalisering av vår næring

Tre eksempler fra et heldigitalt anleggsprosjekt der det offentlige krever leveranse på papir:

**Skiltplaner.** Planer for skilting av vegstrekningen skal godkjennes av Statens vegvesen. Det kreves at disse leveres på papir selv om all informasjon finnes i modellen

**Godkjenning midlertidige tiltak.** Statens Vegvesen godkjenner midlertidige tiltak (omkjøringer etc). Selv om all informasjon er tilgjengelig digitalt, krever godkjenningsmyndigheten dette på papir.

**VA mot kommuner** I et slik vegprosjekt inngår det omfattende VA anlegg. Kommunene krever dette presentert på papir selv om alt er modellert i en digital modell.

#### På dette området foreslår vi:

- Det offentlige må ta en aktiv rolle i digitaliseringen av de verdikjedene der de selv er en part. Vi bidrar gjerne i en diskusjon for å identifisere konkrete punkt i verdikjeden der det offentlige er til hinder for digitalisering.