



Anvendelse av Big Data i organisatorisk styring

En casestudie av Vann- og avløpsetaten i Bergen kommune

Paul Magne Granberg og Henrik Hjelset

Veileder: Andreas Ulfsten

Masterutredning i økonomisk styring

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer innestår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Innholdsfortegnelse

| | |
|---|-----------|
| 1. INNLEDNING | 6 |
| 2. LITTERATUR | 9 |
| 2.1 BIG DATA | 9 |
| 2.1.1 Datadrevne verktøy..... | 12 |
| 2.2 STYRING | 13 |
| 2.2.1 Styringsfunksjoner | 16 |
| 2.3 BIG DATA OG STYRING..... | 19 |
| 2.4 KONSEPTUELT RAMMEVERK | 24 |
| 3. CASEBESKRIVELSE | 25 |
| 3.1 BERGEN KOMMUNE | 25 |
| 3.2 VANN- OG AVLØPSETATEN..... | 26 |
| 3.3 DATASJØEN | 28 |
| 3.4 STYRING I VANN- OG AVLØPSETATEN | 31 |
| 4. METODE | 32 |
| 4.1 VALG AV METODE | 32 |
| 4.1.1 Forskningstilnærming..... | 32 |
| 4.1.2 Forskningsdesign..... | 33 |
| 4.1.3 Forskningsstrategi..... | 33 |
| 4.2 DATAINNSAMLING | 33 |
| 4.2.1 Dokumentanalyse..... | 34 |
| 4.2.2 Utvalg | 34 |
| 4.2.3 Intervjuprosessen..... | 35 |
| 4.2.4 Dataanalyse | 36 |
| 4.3 EVALUERING AV UTREDNINGEN..... | 36 |
| 4.3.1 Prosessuell reliabilitet..... | 36 |
| 4.3.2 Overførbarhet | 37 |
| 4.3.3 Kontekstuell validitet..... | 37 |
| 4.3.4 Selektiv plausibilitet..... | 38 |
| 4.4 ETISKE BETRAKTNINGER..... | 39 |
| 4.5 METODISKE BEGRENSNINGER VED STUDIEN | 39 |
| 5. ANALYSE OG FUNN | 40 |
| 5.1 ANVENDELSE AV BIG DATA TIL OVERVÅKING OG KONTROLL | 40 |

| | |
|--|-----------|
| 5.2 ANVENDELSE AV BIG DATA TIL PLANLEGGING | 51 |
| 5.3 ANVENDELSE AV BIG DATA TIL BESLUTNINGSTAKING | 57 |
| 5.4 ANVENDELSE AV BIG DATA TIL RESULTATMÅLING | 61 |
| 5.5 ANVENDELSE AV BIG DATA TIL RAPPORTERING | 64 |
| 5.6 OPPSUMMERING AV FUNN | 69 |
| 6. DISKUSJON | 70 |
| 6.1 BIG DATA ER 'SYNET' SOM ANVENDES TIL STYRING | 70 |
| 6.2 BIG DATA PÅVIRKER ULIKE DELER AV STYRING OG PÅ ULIKE TIDSPUNKT | 76 |
| 7. KONKLUSJON | 78 |
| 7.1 IMPLIKASJONER OG VIDERE FORSKNING | 79 |
| 8. LITTERATURLISTE | 80 |
| VEDLEGG 1: SEMISTRUKTURERT INTERVJUGUIDE | 84 |
| VEDLEGG 2: INFORMASJONSSKRIV | 85 |
| VEDLEGG 3: SAMTYKKESKJEMA | 86 |
| VEDLEGG 4: NSD SIN VURDERING AV PERSONVERN I PROSJEKTET | 89 |

Forord

Denne masterutredningen er et resultat av det selvstendige arbeidet i hovedprofilen økonomisk styring ved Norges Handelshøyskole. Arbeidet med denne utredningen har vært veldig lærerikt og gitt oss gode erfaringer i å gjennomføre et større vitenskapelig prosjekt.

I denne studien har vi fått mulighet til å forske på et spennende og dagsaktuelt tema; *anvendelse av Big Data i organisatorisk styring*. Gjennom arbeidet med denne oppgaven har vi fått nye og verdifulle innsikter som kan bidra til å forstå samspillet mellom disse to konseptene. Det har vært spesielt interessant å forske på dette i en bestemt kontekst, nemlig Vann- og avløpsetaten i Bergen kommune, noe som har gitt oss konkrete og dype innsikter i en spennende bransje.

Vi ønsker å rette en stor takk til Hildegunn Grindheim, Ali Mekki og Kjersti Steindal i Bergen kommune for gode diskusjoner, verdifulle innspill og imøtekommende holdninger. Deres bidrag til valg av case og vinkling av utredningen har vært til stor hjelp. Videre ønsker vi å takke alle våre informanter i Vann- og avløpsetaten i Bergen kommune og Bergen Vann. Takk for engasjerende og spennende samtaler, og for interessante innsikter i deres arbeid.

Til slutt ønsker vi å takke vår veileder, Andreas Ulfsten, for god veiledning, solid oppfølging og verdifulle tilbakemeldinger underveis i prosessen. Dine tanker og innspill har vært avgjørende for utformingen av denne oppgaven, og har hjulpet oss å forstå hva som er relevant å studere i et akademisk perspektiv.

Vi ønsker deg god lesning!

Bergen, 31 mai 2021

Paul Magne Granberg & Henrik Hjelset

Sammendrag

Formålet med denne masterutredningen har vært å undersøke hvordan Big Data anvendes i organisatorisk styring. Fremveksten av stadig nye digitale enheter, medfører at det produseres mer data nå enn noen gang tidligere, et fenomen som gjerne omtales Big Data. Gitt at Big Data kan anses som en kilde til informasjon, er dette et fenomen som potensielt kan ha store innvirkninger på organisatorisk styring. På tross av dette, er det få empiriske studier som har studert sammenhengen mellom Big Data og styring. Vi har derfor ønsket å belyse følgende problemstilling: “*Hvordan anvendes Big Data i organisatorisk styring?*”

Ettersom vi var interessert i å studere *hvordan* Big Data anvendes i styring, valgte vi å gjennomføre en eksplorativ studie med kvalitative data. For å studere fenomenet i en organisatorisk kontekst valgte vi en casestudie av Vann- og avløpsetaten i Bergen kommune. Dette gav oss muligheten til å studere detaljer og nyanser om fenomenet i en gitt kontekst. Ettersom Vann- og avløpsetaten i Bergen kommune er offentlig virksomhet med store ambisjoner om å utnytte potensialet ved Big Data, anså vi dette å være et spennende og relevant case. Datagrunnlaget i studien er hovedsakelig basert på semi-strukturerte dybdeintervju med ansatte i Vann- og avløpsetaten.

Våre funn indikerer at Big Data har innvirkning på styring, men (i) ulike deler av styring og (ii) på ulike tidspunkt. Big Data virker å gi en umiddelbar gevinst til overvåking og kontroll ved at man får bedre oversikt, særlig i en bransje hvor avstanden mellom mennesker og kapital er stor. Samtidig medfører økt bruk av eksterne data at enkelte styringsfunksjoner, slik som planlegging og beslutningstaking, kan bli betydelig mer komplekse. Dette krever ny kompetanse og bruk av avanserte analyseverktøy, noe som kan føre til at det tar noe lengre tid før man kan realisere gevinster fra bruk av Big Data i disse funksjonene.

1. Innledning

Denne masteroppgaven vil ta for seg hvordan Big Data anvendes i organisatorisk styring. Big Data er et fenomen som omhandler informasjon fra enorme datamengder, og kan ha store innvirkninger på styring. Tilgang på informasjon fra store datamengder kan blant annet bidra til et bedre beslutningsgrunnlag, oppdagelse av nye innsikter, gi dypere kundeforståelse og kan føre til automatisering av forretningsprosesser (McAfee, Brynjolfsson, Davenport, Patil & Barton, 2012; Davenport, Bart & Bean, 2012). Bruk av Big Data har potensiale til å kunne påvirke hvordan organisasjoner jobber med sentrale styringsfunksjoner som overvåking, planlegging, beslutningstaking og rapportering. Ved å benytte Big Data som grunnlag for analyser og beslutninger kan organisasjoner bevege seg mot det som ofte refereres til som 'evidensbasert beslutningstaking' (Holsapple, Lee-Post & Pakath, 2014), der data er hovedgrunnlaget for beslutninger fremfor subjektiv erfaring og ekspertise.

Til enhver tid genereres det enorme mengder data. Bare i løpet av perioden fra 2016 til 2018, ble 90 % av all tilgjengelig data i verden produsert (Marr, 2018). Den kontinuerlige utviklingen av teknologier som Internet of Things, sosiale medier og sensorer gjør at data i større grad kan genereres i sanntid. Dette medfører at vi stadig får tilgang til mer data. Til tross for større tilgang på data, gjenstår det fortsatt ubesvarte spørsmål knyttet til *hvordan* organisasjoner kan benytte slike store datamengder til styring på en fornuftig måte (Kitchin, 2014; Grover, Chiang, Liang & Zhang, 2018).

Styring er et bredt begrep, og har i økonomisk sammenheng blitt mye studert, blant annet under begrepet *Management Accounting*. I korte trekk kan Management Accounting forstås som en funksjon som skal forsyne ledere og ansatte i en organisasjon med relevant finansiell og operasjonell informasjon, slik at disse personene kan fatte gode beslutninger, allokere ressurser, samt overvåke og evaluere organisasjonens ytelse (Atkinson, Kaplan, Young, 2011). Gitt at Big Data kan anses som en kilde til informasjon, og styring er en funksjon som benytter informasjon, er det en naturlig sammenheng mellom disse konseptene. Relasjonen mellom informasjonsteknologi og styring har lenge vært et område av stor interesse for styringslitteraturen, spesielt fokuset på teknologiens innvirkning på styring (Knudsen, 2020). På tross av dette, hevdes det at tidligere forskning har fokusert på det som nå er utdaterte teknologier (Granlund, 2011), og at ettersom teknologi er et dynamisk fenomen trengs det nyere forskning på teknologiens innvirkning på styring (Prasad & Green, 2015).

Big Data er et relativt nytt fenomen innen informasjonsteknologi, og er derav lite studert innenfor styringslitteraturen. Den generelle oppfattelsen blant forskere er at Big Data vil være et disruptivt og revolusjonerende fenomen innen styring (Rikardson & Yigitbasioglu, 2018; McAfee et al., 2012). Det store spørsmålet knyttet til Big Data, har derimot vært hva man skal gjøre med de store datamengdene og *hvordan* organisasjoner kan dra nytte av dem til å skape verdi for virksomheten.

I denne masteroppgaven ønsker vi å bidra til den akademiske debatten rundt Big Data's innvirkning på styringsfunksjoner. Rikardson & Yigitbasioglu (2018) har i sin litteraturgjennomgang identifisert fire forskningsartikler som omhandler dette temaet. Alle de fire artiklene var konseptuelle, og ingen av dem var basert på empirisk forskning. Forfatterne av denne litteraturgjennomgangen hevder derfor at det trengs flere kvalitative casestudier for å demonstrere *hvordan* Big Data brukes til organisatorisk styring.

For å undersøke dette fenomenet nærmere, ønsker vi å belyse følgende problemstilling:

“Hvordan anvendes Big Data i organisatorisk styring?”

Ettersom vi er interessert i å studere *hvordan* Big Data anvendes i styring, har vi valgt en eksplorativ studie med kvalitative data. For å studere fenomenet i en organisatorisk kontekst har vi valgt en casestudie av Vann- og avløpsetaten (VA-etaten) i Bergen kommune. Dette har bidratt til å få frem detaljer og nyanser om fenomenet. Datagrunnlaget i studien er hovedsakelig basert på semi-strukturerte dybdeintervju med ansatte i VA-etaten. Bergen kommune har i flere år satset tungt på digitalisering, og lanserte i 2019 sin datasjø – ‘Lungegårdsvannet’. Datasjøen er en metodikk for innsamling, lagring og tilgjengeliggjøring av store og varierte datamengder (Bergen kommune, u.d.-b), altså et lokalt og konkret uttrykk for konseptet Big Data. Formålet med datasjøen til Bergen kommune er tilgjengeliggjøring av data for både interne og eksterne aktører, tilrettelegging for innovasjon, og bedre innsikt og beslutningsstøtte gjennom bruk av data (Bergen kommune, u.d.-b). Bergen kommune virker med dette å ha et sterkt ønske om å benytte Big Data for å tilby bedre tjenester for innbyggerne.

VA-etaten er en av de fremste brukerne av kommunens datasjø. Via datasjøen henter de blant annet måledata hvert minutt fra 13.000 sensorer som er plassert rundt om på vann- og avløpsnettet (Mekki, 2020). Formålet med bruk av sanntidsdata er at det skal bidra med bedre

oversikt over vann- og avløpsnettet, samt gi bedre beslutningsgrunnlag for diverse driftstiltak. Etersom VA-etaten i Bergen kommune er offentlig virksomhet med store ambisjoner om å utnytte potensialet ved Big Data, anser vi dette å være et spennende og relevant case for å undersøke problemstillingen.

2. Litteratur

I teoridelen skal vi presentere to konsepter som vi benytter videre i oppgaven, nemlig Big Data og styringsfunksjoner. Her vil vi definere konseptene og gjennomgå tidligere litteratur. Til slutt vil vi se på sammenhengen mellom de to konseptene og trekke inn eksisterende litteratur som etterspør mer forskning på dette området.

2.1 Big Data

Fremveksten av nye teknologier og digitale enheter fører til at stadig mer data genereres. I dagens samfunn er det nærmest ingen områder som er uberørt av digitale teknologier og vi er stadig mer koblet til nettet (Bhimani & Willcocks, 2014). Dette medfører at de fleste handlingene vi gjør etterlater seg digitale spor, slik som e-postmeldinger, korttransaksjoner i en dagligvarebutikk, klikk-stream på internett, bilder og video fra sosiale medier, og sensorregistrering fra smarte enheter (Zhang, Yang & Appelbaum, 2015). Denne utviklingen hvor det produseres nye og store mengder data har ført til at det har vokst frem et fenomen som gjerne omtales *Big Data*.

Big Data er et relativt nytt fenomen og det er derfor gjort begrenset med studier for å forstå dette fenomenet. Det er et noe uklart begrep, og det eksisterer foreløpig ingen etablert felles definisjon, noe som kan være grunnen til at forskningsartikler definerer det noe ulikt. En utbredt forståelse definerer imidlertid Big Data som datasett så store eller ustrukturerte at de ikke enkelt kan prosesseres eller analyseres ved hjelp av tradisjonelle databaser eller andre systemer (Warren, Moffitt & Byrnes, 2015; McAfee et al., 2012). Med andre ord virker omfanget av Big Data å være så stort at det krever nye strukturer og metoder for å lagre og analysere de enorme datamengdene. En videre utbredt forståelse av begrepet kategoriserer Big Data med 4 V'er; Volume, Velocity, Variety og Veracity (Laney, 2001; Gandomi & Haider, 2015). Disse V'ene fungerer som kjennetegn på Big Data; et enormt volum, høy hastighet, bred variasjon, og usikker kvalitet (Appelbaum, Kogan, Vasarhelyi & Yan, 2017). For å gi oss en bedre forståelse for omfanget av Big Data, skal vi se nærmere på de fire dimensjonene.

Volum referer til den enorme mengden data som en organisasjon eller individer genererer (Lee, 2017; Simon, 2013). Den globale datamengden hatt en kraftig utvikling de siste årene (Reinsel, Gantz & Rydning, 2017), og ifølge Marr (2018) ble hele 90 % av all tilgjengelig data i verden

per 2018 produsert kun i løpet av de siste to årene. Dette er altså en eksponentiell utvikling som også ventes å fortsette i årene fremover (Reinsel et al, 2017).

Denne veksten er i hovedsak drevet av at nye digitale enheter er koblet til nettet (Lee, 2017). For å gi et bilde av hvor enorme datamengder det er snakk om, produseres det ifølge Marr (2018) rundt 2,5 exabytes, eller 2,5 billioner gigabytes med data hver eneste dag. Det er også estimert at den amerikanske dagligvarekjeden Walmart alene samler inn mer enn 2,5 petabytes, eller 2,5 millioner gigabytes med data av forskjellige karakter hver dag (McAfee et al., 2012). Lee (2017) hevder at en datamengde må være på minst 1 terabyte, eller 1000 gigabyte, for at den skal kunne omtales som Big Data. Samtidig bemerker han at denne grensen stadig flyttes etterhvert som teknologien utvikler seg. Andre har påstått at kriteriet om volum er møtt når datasettet er så stort at man ikke kan samle, lagre og analysere det med tradisjonelle data- og statistiske metoder (Holmes, 2017). Uansett får vi her en forståelse av at Big Data handler om enormt store volum og data som er vanskelig å håndtere for organisasjoner.

Velocity referer til hastigheten som data genereres og prosesseres i (Lee, 2017). Big Data genereres ofte i sanntid eller tilnærmet sanntid, og den registreres kontinuerlig fra ulike kilder som internett, smarttelefoner, GPS og sensorer (Rikhardson & Yigitbasioglu, 2018; Kitchin, 2014). Sanntidsdata gjør det mulig for en virksomhet å være mer tilpassningsdyktig, og er av den grunn ansett å være en veldig verdifull egenskap ved Big Data (McAfee et al., 2012). Et eksempel på sanntidsdata er selvkjørende biler som er avhengige av kontinuerlig sensordata og analyse av denne for å kunne være pålitelig (Holmes, 2017). Velocity har også sammenheng med volum, hvor kontinuerlig registrering og lagring av data følgelig gir et enormt volum over tid. Et eksempel på dette finner man hos bilprodusenten Tesla som per august 2020 hadde samlet inn sensordata fra nesten 5 milliarder kjørte kilometer fra sine biler (Lambert, 2020). Dette kan igjen brukes til å forbedre den selvkjørende teknologien til selskapet. Flere og flere selskaper ser nå et økende behov for sanntidsdata og det er forventet at dette vil være en viktig driver for den videre globale dataveksten (Reinsel et al, 2017).

Variety refererer til variasjon i datatypene som produseres (Simon, 2013). Big Data kan oppstå fra mange ulike kilder, både interne transaksjonssystemer og databaser, samt fra eksterne kilder som sosiale medier, værmeldinger og IoT enheter (Zhang et al., 2015). I tillegg til variasjon fra interne og eksterne datakilder, skiller man gjerne også mellom variasjon av strukturerte og ustrukturerte data. Strukturerte data, som navnet tilsier, har en strukturert og

standardisert form. Den lagres ofte i databaser eller regneark i form av tabeller, og er derav mer forutsigbar, gjenkjennelige og enkle å behandle (Marr, 2018). Strukturert data er oftest generert internt i virksomheten, slik som transaksjonsdata, varelager og kundeundersøkelser (Lee, 2017). Fremveksten av nye digitale enheter gjør imidlertid at mer av den tilgjengelige dataen er såkalt ustrukturert data slik som e-poster, lydfiler, sensorregistrering og forskjellig aktivitet i sosiale medier (Appelbaum et al., 2017). Denne ustrukturerte dataen er vanskeligere å analysere grunnet sin ustrukturerte natur og krever ofte bruk av mer avansert analyseverktøy som Big Data Analytics (Holmes, 2017). Gandomi & Haider (2015) hevder at så mye som 95% av all data er ustrukturert. Dette byr følgelig på en stor utfordring for virksomheter som ønsker å dra nytte av Big Data.

Den siste V'en, *Veracity*, handler om kvaliteten på dataen. For at dataen skal være nyttig i et analyseformål er det viktig at dataen er av høy kvalitet og kan stoles på (Chae, Yang, Olson & Sheu, 2014). Redman (2013) beskriver data med høy kvalitet som komplett, presis, gyldig, relevant og tidsaktuell. Selv om det genereres mer data er det ikke nødvendigvis slik at all denne dataen er nyttig og gir verdi til virksomheten. Etersom Big Data består av mange forskjellige datatyper, er det gjerne større risiko og usikkerhet knyttet til kvaliteten på data. Ustrukturerte data fra for eksempel sosiale medier kan ofte være upresise og inneholder gjerne mye subjektivitet, noe som kan gi en usikkerhet med tanke på påliteligheten til denne type data (Holmes, 2017; Lee, 2017). Usikker datakvalitet kan derfor anses som en utfordring ved bruk av Big Data.

Andre har forsøkt å utvide denne definisjonen ved å legge til flere dimensjoner. Det er blant annet foreslått dimensjoner som omhandler gyldighet, volatilitet og verdi (Eilars, 2013). Selv om flere dimensjoner på den ene siden kan bidra til å forstå begrepet bedre, kan det på den andre siden medføre mer kompleksitet og usikkerhet knyttet til hva som kan forstås som Big Data. I denne oppgaven anser vi at de fire V'ene Volume, Velocity, Variety og Veracity gir et tydelig nok bilde til å forstå hva Big Data kan være. Vi inkluderer derfor ikke ytterligere dimensjoner av begrepet i denne oppgaven.

Oppsummert gir de fire V'ene et godt bilde av mulighetene, men også kompleksiteten ved Big Data. Tilgang til enorme datamengder på detaljnivå gjør at virksomheter har mulighet til å hente ut verdifulle innsikter og sammenhenger som de tidligere ikke har kunnet gjøre med aggregert informasjon fra tradisjonelle databaser. Samtidig kan det enorme volumet og den

store variasjonen i datatyper medføre at det for mange oppleves overveldende å jobbe med slike store datasett.

2.1.1 Datadrevne verktøy

I seg selv er ikke Big Data noe mer enn enorme mengder forskjellig typer data. For å utnytte potensiale til disse datamengdene, må den anvendes og tas i bruk gjennom verktøy og metoder som kan hente ut verdifulle innsikter. Datadrevne verktøy er én måte å anvende de store datamengdene på. I denne delen vil det gis en kort introduksjon og beskrivelse av noen generelle kategorier med slike verktøy. Dette er verktøy som er relevante i case-organisasjonen og som kan være nyttig for leseren å kjenne til ved senere analyse og diskusjon i oppgaven.

Datavisualiseringsverktøy

Datavisualiseringsverktøy er blitt et utbredt fenomen i organisasjoner (Appelbaum et al., 2017). Ledere og ansatte i bedrifter forventer i større grad at analyser og målinger fremstilles i et format som er lett å forstå og som kan gi raske innsikter i ulike situasjoner (Appelbaum et al., 2017). Davenport og Kim (2013), hevder at visualiseringer hjelper folk å forstå data gjennom en visuell fortelling av hendelser fra start til slutt.

Det finnes mange forskjellige måter å visualisere data på, for eksempel gjennom karttjenester, dashboards, grafer og figurer. Eksempler på populære visualiseringsverktøy er blant annet Microsoft Power-BI og Tableau. Styringsfunksjonen forsyner informasjon til beslutningstakere, og det forventes dermed at den klarer å presentere informasjon på en effektiv og klar måte. Bruk av visualiseringsverktøy vil kunne være en stor hjelp til dette (Appelbaum et al., 2017).

Maskinlæring

Maskinlæring er et felt innenfor det større begrepet *kunstig intelligens*, hvor man bruker matematiske og statistiske metoder for å la datamaskinen finne mønster og sammenheng i store datamengder (Tidemann & Elster, 2019). Ved hjelp av probabilistiske metoder kan datamaskinen identifisere mønstre i datamengden, som den deretter bruker til å lage en datamodell som kan gjøre prediksjoner (Moll & Yigitbasioglu, 2019). Etter hvert som datamodellen får flere datamengder og mer erfaring, blir resultatene mer nøyaktige.

Datamaskinen vil dermed 'lære' og få erfaring omtrent på samme måte som et menneske (Microsoft, u.d.).

Maskinlæring er ansett å være en av de store teknologiene for å anvende Big Data (Moll & Yigitbasioglu, 2019). Big Data kjennetegnes som nevnt av stor variasjon i type data, og mye av dette oppstår fra eksterne kilder utenfor organisasjonen. Tilgangen til slike data gjør at man kan sammenstille data fra ulike kilder og derav søke etter nye sammenhenger og innsikter fra et større datagrunnlag. Dette kan for eksempel bidra til å forklare variasjoner i ulike driftsprosesser som man tidligere ikke har kunnet gjøre basert på kun interne datakilder. På en annen side blir gjerne sammenstillingen av disse datasettene enormt store og for komplekse til at det menneskelige sinn klarer å se sammenhenger. Det er her mange har forhåpninger til at maskinlæring kan benyttes som et verktøy med kapasitet til å kunne hente ut innsikt fra datamengdene (Moll & Yigitbasioglu, 2019).

2.2 Styring

Styring er et bredt begrep, og kan konseptualiseres på mange måter. I økonomisk sammenheng har styring blitt mye studert, blant annet under begrepet *Management Accounting (MA)*. Ettersom vi er interessert i å forske på anvendelsen av informasjonsteknologi innenfor økonomisk styring, benytter vi MA-litteraturen for å konseptualisere styring i denne oppgaven. Det kan derfor være hensiktsmessig først å se nærmere på utviklingen av MA-begrepet, før vi senere forklarer hvordan vi velger å konseptualisere styring.

Gjennom de siste årene har rollen til såkalte *management accountants (MAs)* - personer som i det daglige arbeider med økonomisk styring - endret seg betraktelig (Windeck, Weber & Strauss, 2013). Tradisjonelt sett var hovedoppgaven for MAs å fremskaffe intern finansiell informasjon om tidligere hendelser, samt å vedlikeholde regnskapssystemer (Windeck et al., 2013). Dette kan i stor grad minne om jobben til en tradisjonell regnskapsfører som driver intern kontroll av historisk finansiell informasjon. Disse arbeidsoppgavene gjorde at MAs ofte ble omtalt som *bean counters* eller *score keepers* ettersom de i stor grad kun rapporterte informasjon om historiske finansielle verdier (Windeck et al., 2013).

Senere kom det derimot kritikk om at funksjonen som en *bean counter* var for snever og for lite relevant, og flere mente at MA-funksjonen måtte innebære mer enn bare finansiell rapportering av historiske hendelser (IMA, 2008). Johnson og Kaplan (1987) kritiserte den tradisjonelle formen for MA og hevdet blant annet at informasjonen som ble rapportert var for lite tidsrelevant, for aggregert og for uklar til å være verdifull for beslutningstakere. De hevdet dermed at MAs jobbet på en for reaktiv måte, med fokus kun på deskriptiv finansiell analyse av tidligere hendelser. Andre har også senere påpekt at dagens forretningsverden krever tidsrelevant data, og dermed er ikke historiske aggregerte måltall en ideell kilde til et godt beslutningsgrunnlag, ettersom de kun rapporterer hva som *har* skjedd (Nielsen, 2018). Flere mente derfor at MAs måtte være involvert i mer strategiske og fremoverrettede aktiviteter (IMA, 2008).

På bakgrunn av denne kritikken har man i senere tid sett et skifte hvor MAs går fra *bean counters* til *business partners* (Windeck et al., 2015). Denne overgangen innebærer at MAs har et mer strategisk fokus hvor man er aktivt involvert i ulike typer forretningsprosesser, utøver kontroll og på denne måten støtter oppunder ledere som tar beslutninger (Windeck et al., 2013). Til forskjell fra den tradisjonelle formen for MA har man altså her høyere grad av interaksjon med diverse organisatoriske funksjoner og beslutningstakere (Windeck et al., 2013). Utviklingen av informasjonsteknologi virker å være en sentral driver for denne utviklingen. Rutinemessige arbeidsoppgaver som innsamling og rapportering av historisk informasjon har blitt en mindre del av rollen, ettersom informasjonsteknologier som ERP-systemer har automatisert mye av dette (IMA, 2008).

En ytterligere konsekvens av den teknologiske utviklingen er at denne informasjonen også blir mer tilgjengelig for andre faggrupper i organisasjonen. Tradisjonelt har forskning på MA i stor grad omtalt økonomen i den sentrale rollen som Management Accountant. Dette kan henge sammen med at tidligere forståelser av MA-begrepet innebar oppgaver med fokus kun på finansiell informasjon, slik som innhenting og analyse av regnskapstall, og presentering av finansielle rapporter til eiere, kreditorer og myndigheter (IMA, 2008). Ettersom økonomer gjerne er best egnet til å utføre slike finansielle oppgaver, er det forståelig at økonomens rolle i MA har vært et utbredt fokus. Senere perspektiver på MA fokuserer derimot også på ikke-finansielle aspekter (Atkinson, Kaplan & Young, 2011; Appelbaum et al., 2017), som for eksempel kan innebære vurdering av operasjonelle måltall.

Dette kan tyde på at MA-litteraturen har blitt bredere i sitt fokus, en utvikling som også belyses av Caglio (2003) som i sin artikkel beskriver hvordan tilgangen på nye ERP-systemer utfordrer den tradisjonelle regnskaps (*accounting*) -rollen. Han introduserer her begrepet *hybridization* som en referanse til at andre faggrupper i organisasjonen nå kan utføre oppgavene som tidligere var forbeholdt økonomene ettersom utviklingen av nye teknologiske løsninger har ført til bedre tilgjengelig informasjon (Caglio, 2003). Enkelte studier hevder at det er tendenser til uklare grenser for MA-profesjonen, da digitalisering gjør at den organisatoriske posisjonen til tradisjonelle MAs utfordres av andre profesjonelle grupper (Knudsen, 2020). Bruken av tittelen «Management accountant» har også blitt drastisk redusert de siste tiårene (Otley, 2008), noe som kan tale for at det i dag er få som jobber *kun* i denne rollen.

Disse uklare grensene kan også være en av grunnene til hvorfor mange i dag mener vi har et tydelig skille mellom MA som profesjon, og MA som en prosess eller som et sett av oppgaver som må utføres av personer i organisasjonen (Rom og Rohde, 2007). Dette skillet kan man også se reflekteres i forskjellige definisjoner av begrepet. Institute of Management Accountants gir følgende definisjon av MA:

«(...) a *profession* that involves partnering in management decision making, devising planning and performance management systems, and providing expertise in financial reporting and control to assist management in the formulation and implementation of an organization's strategy». (IMA, 2008)

Atkinson, Kaplan og Young (2011), definerer imidlertid MA som: «(...) the *process* of supplying the managers and employees in an organization with relevant information, both financial and nonfinancial, for making decisions, allocating resources, and monitoring, evaluating and rewarding performance».

Fra disse definisjonene ser vi at finnes det flere ulike forståelser av hva MA er, men det er samtidig klare likhetstrekk. Det klareste fellestrekket virker å være at funksjonen skal forsyne ledere og ansatte i en organisasjon med relevant informasjon slik at disse personene kan fatte gode beslutninger. Den største forskjellen i disse definisjonene er imidlertid om MA forstås som en rolle eller en prosess. De konkrete oppgavene som definisjonene refererer til, skal vi se nærmere på i neste del av teorien om styringsfunksjoner.

På tross av uklare grenser for selve Management Accountant-profesjonen, virker de konkrete MA-funksjonene å være en sentral del for å forstå økonomisk styring. Utviklingen av informasjonsteknologi har ført til at andre faggrupper i organisasjonen også driver med styring, ikke bare den konkrete rollen Management Accountant. Vår konseptualisering av styring i denne oppgaven innebærer derfor fokuset på de konkrete MA-funksjonene som utføres av forskjellige personer i organisasjonen. I neste del skal vi derfor se nærmere på hva disse konkrete funksjonene innebærer.

2.2.1 Styringsfunksjoner

Innenfor litteraturen om funksjonen økonomisk styring finnes det flere aktuelle forskningsområder. Litteraturen har definert fire hovedelementer innenfor MA som fungerer som en overordnet struktur for å forstå konseptet (Rom & Rhode, 2007; Knudsen, 2020; Rikardson & Yigitbasioglu, 2018). Det første elementet handler om de overordnede styringsfunksjonene eller oppgavene til MA-funksjonen slik som planlegging, rapportering og beslutningstaking. Det andre elementet handler om de mer spesifikke MA-teknikkene og metodene som brukes til å gjennomføre oppgavene slik som aktivitetsbasert kostnadsføring, strategikart og prognoser. Det tredje elementet omhandler organiseringen av MA-funksjonen, herunder hvordan funksjonen er organisert på tvers av ulike industrier og organisasjonsstrukturer. Det fjerde og siste elementet handler om MAs innvirkning på organisasjonen og de ansattes atferd. Etersom vi ønsker å studere hvordan informasjonsteknologi anvendes i organisatorisk styring, har vi i denne studien valgt å fokusere på det første hovedelementet som omhandler de overordnede styringsfunksjonene til MA-funksjonen.

Styringsfunksjoner ser vi i lys av det Management Accounting-litteraturen ser på som sentrale oppgaver for styring (Rikardsson & Yigitbasioglu, 2018; Rom & Rhode, 2007; Knudsen, 2020). Litteraturen påpeker at disse oppgavene kan variere ut fra kontekst og situasjon i organisasjonen, noe som kan være grunnen til at det er noe variasjon i hva de ulike forskningsartikler belyser som sentrale MA-oppgaver. Det kan for eksempel tenkes at en ideell organisasjon har mest fokus på å holde kostnadsnivået lavt, mens en profittmaksimerende organisasjon i tillegg til kostnader også har stort fokus på styring av inntekter. Videre kan det også tenkes at en driftsorganisasjon har et mer utbredt fokus på de operasjonelle funksjonene,

mens en bank naturligvis også har stort fokus på de finansielle funksjonene. Dette vil følgelig kunne påvirke hvilke styringsfunksjoner som er mest utbredt i den enkelte organisasjonen. En annen grunn til at oppgavene fremstilles noe ulikt kan være at det er vanskelig å sette tydelige grenser mellom de ulike funksjonene, spesielt ettersom konteksten mellom ulike organisasjoner kan variere stort.

Ifølge litteraturgjennomgangen til Rikhardsson og Yigitbasioglu (2018) innebærer styringsfunksjonen oppgaver som planlegging, kontroll, resultatmåling, transaksjonsanalyse, rapportering og beslutningstaking. Siden dette er en litteraturgjennomgang er disse oppgavene en oppsummering av hva flere tidligere forskningsartikler fremhever som sentrale oppgaver for funksjonen. En annen litteraturgjennomgang, Knudsen (2020), belyser transaksjonsanalyse, rapportering og beslutningstaking som sentrale oppgaver. Denne artikkelen benytter de samme tre oppgavene som Rom & Rhode (2007) benytter i sin litteraturgjennomgang. Videre klassifiserer Appelbaum et al. (2017) i sin artikkel funksjonen inn i kostnadsføring, resultatmåling, planlegging og beslutningstaking. Fra MA-definisjonen til Atkinson et al. (2011) beskrives også noen sentrale styringsfunksjoner, slik som beslutningstaking, fordeling av ressurser, og overvåking og evaluering av resultater.

Vi ser altså fra disse forskningsartiklene at flere styringsfunksjoner virker å være sentrale. For det første er beslutningstaking en sentral oppgave som fremheves i alle de ovennevnte artiklene. Videre er rapportering, planlegging, transaksjonsbehandling og resultatmåling tre oppgaver som er felles i minst to av artiklene. Deretter er det flere oppgaver som ikke fremheves i samme grad, og er derav mer unik for den enkelte artikkel. På en annen side er flere av disse oppgavene veldig like av karakter. Kontroll og overvåking er for eksempel lignende konsepter, og kan derfor tolkes å innebære overlappende praksiser. Samtidig kan planlegging og fordeling av ressurser sees på som nokså like oppgaver hvor fordeling av ressurser gjerne krever planlegging for å kunne gjennomføres på en god måte. Resultatmåling og evaluering kan også argumenteres å henge nøye sammen, ettersom måling av resultater gjerne gjennomføres for å evaluere hvordan organisasjonen presterer på forskjellige måltall.

Sammenhengen mellom flere av styringsfunksjonene kan noen ganger gjøre det utfordrende å skille dem fra hverandre. Overvåking og kontroll av driften henger for eksempel tett sammen med planlegging av driftstiltak. Basert på informasjon som samles inn gjennom overvåking, gjør man vurderinger og planlegger hvilke driftstiltak man skal sette inn. Det samme gjelder

rapportering, spesielt internt, hvor man gjerne agerer og gjør noen beslutninger basert på hva som forekommer i rapporten. Resultatmåling er også en funksjon som kan henge tett sammen med rapportering, hvor måltallene i resultatmåling gjerne brukes til rapportering. Til slutt er beslutningstaking en funksjon som gjerne blir utført som en konsekvens av de andre funksjonene. Basert på informasjon, agerer man og gjør ulike beslutninger.

Når vi snakker om styringsfunksjoner i oppgaven, referer vi altså til de konkrete MA-funksjonene eller oppgavene som utføres av ulike personer i organisasjonen. Grunnet omfanget på denne masteroppgaven, vil det være lite hensiktsmessig å gjennomgå alle de ulike styringsoppgavene som litteraturen fremlegger. Hvilke konkrete styringsfunksjoner vi skal se på i denne oppgaven er derfor basert på en kombinasjon av hva tidligere litteratur sier er sentralt, og hvilke styringsfunksjoner vi opplever er passende i konteksten av case-organisasjonen i denne studien. Av den grunn skal vi fokusere på de operasjonelle funksjonene overvåking og kontroll, planlegging, beslutningstaking, resultatmåling og rapportering. Under casebeskrivelsen i kapittel 3, vil vi redegjøre nærmere for disse styringsfunksjonene sett opp mot caset i Bergen kommune.

For å oppsummere dette kapittelet om styring og styringsfunksjoner, har vi sett at styring er et konsept som kan forstås på flere ulike måter. En felles forståelse av styringsfunksjonen virker derimot å være at funksjonen skal frembringe informasjon for beslutningstakere. Dette vil også være et sentralt aspekt ved vår konseptualisering av styring for denne oppgaven. Videre har vi sett at styring kan studeres med flere ulike innfallsvinkler, blant annet ved å se på organisering av funksjonen, de konkrete styringsfunksjonene, og hvordan funksjonen påvirker atferden til de ansatte. Ettersom vi ønsker å se på anvendelse av informasjonsteknologi i organisasjonens styringsfunksjon, har vi valgt å konseptualisere styring i denne oppgaven som de konkrete styringsfunksjonene.

2.3 Big Data og styring

Gitt at Big Data kan anses som en kilde til informasjon, og styring kan defineres som en funksjon som skal frembringe og bruke informasjon til beslutningstaking, er det en naturlig sammenheng mellom disse fenomenene. Relasjonen mellom informasjonsteknologi og styring har lenge vært et område av stor interesse for styringslitteraturen, spesielt fokuset på teknologiens innvirkning på styring (Knudsen, 2020; Rom & Rhode, 2007). På tross av dette, hevdes det at tidligere forskning har fokusert på det som nå er utdaterte teknologier (Granolund, 2011), og at ettersom teknologi er et dynamisk fenomen trengs det nyere forskning på teknologiens innvirkning på styring (Prasad & Green, 2015).

Big Data er en av disse nye teknologiene, men få studier sier noe om innvirkningen Big Data har på styring. I sin litteraturgjennomgang har Rikhardson & Yigitbasioglu (2018) kun identifisert fire artikler som ser på sammenhengen, men alle disse er konseptuelle studier. Forfatterne etterspør derfor mer empirisk forskning som studerer hvordan Big Data brukes til styring, spesielt innenfor ulike styringsfunksjoner. I de neste avsnittene vil vi belyse hva eksisterende litteratur sier om sammenhengen mellom begrepene, samt diskutere hvilke muligheter og utfordringer Big Data gir for styringsfunksjonen.

På tross av lite forskning, hevder flere at Big Data vil være en disruptiv kraft innenfor styring (Rikhardson & Yigitbasioglu, 2018). Tilgang på detaljert informasjon fra store datamengder kan blant annet kan gi bedre beslutningsgrunnlag, oppdagelse av nye innsikter, dypere kundeforståelse, og føre til automatisering av forretningsprosesser (McAfee et al., 2012; Davenport et al, 2012). Et sentralt fokus i litteraturen har handlet om hvordan Big Data vil endre Management Accountant-rollen. Her hevdes det blant annet at rollens oppgave med å registrere data vil bli mindre viktig (Rikhardson & Yigitbasioglu, 2018), noe som er i tråd med utviklingen av Management Accountant-rollen som vi har sett på tidligere. Videre hevdes det at folk som driver med styring trenger nye ferdigheter for å jobbe med Big Data, spesielt kompetanse på å jobbe med statistiske verktøy (Rikhardson & Yigitbasioglu, 2018).

Styring skal forsyne beslutningstakere med relevant informasjon, og kan derfor dra stor nytte av Big Data. Her legger litteraturen vekt på særlig fire kjennetegn ved Big Data som kan skape store muligheter (Rikhardson & Yigitbasioglu, 2018). Den første er at Big Data innebærer konstant registrering av sanntidsdata. For det andre innebærer Big Data registrering av både

intern og ekstern data. Den tredje er at Big Data gir mulighet til å jobbe med data om hele populasjoner, istedenfor kun å benytte utdrag. Den siste handler om at Big Data gir mulighet til å benytte ustrukturerte datakilder til analyse.

For det første innebærer Big Data konstant registrering av sanntidsdata (Rikhardson & Yigitbasioglu, 2018). Dette gir styringsfunksjonen tidsrelevant informasjon som gir mulighet for raske innsikter til enhver tid. Styringsfunksjoner som overvåking og kontroll kan dermed få et raskt og oppdatert informasjonsgrunnlag til å vurdere tilstanden på ulike forretningsprosesser. Videre innebærer Big Data registrering av både intern og ekstern data (Rikhardson & Yigitbasioglu, 2018). Dette gir muligheten til å sammenstille data fra flere ulike kilder som sammen kan brukes til å søke etter nye trender og sammenhenger i datamaterialet (Constantiou & Kallinkikos, 2015). Det at man tar i bruk data fra andre kilder enn interne systemer kan også føre til at man får bedre prognoser og estimater av ressursbruk (Warren et al., 2015). For eksempel kan et busselskap sammenstille intern data om bussruter med ekstern data om trafikkflyt for å gi et bedre estimat på hvor lang tid bussen kommer til å bruke på ruten til forskjellige tider av døgnet. Dette vil være verdifull informasjon som for eksempel kan bidra til bedre planlegging og allokering av ressurser.

Videre påpekes det at Big Data gir mulighet til å jobbe med hele populasjoner istedenfor kun å benytte utdrag (Rikhardson & Yigitbasioglu, 2018). Big Data søker etter å fange opp og forstå hele populasjoner hvor $n = \text{alle}$ (Kitchin, 2014). Hvor et utdrag av populasjonen kun gir en antagelse av hva som gjelder, vil Big Data gi muligheten til å forstå hva som faktisk gjelder for hele populasjonen. Dette vil for eksempel kunne være nyttig for å tilpasse annonsering og reklame til den enkelte kunde basert på deres aktiviteter. Til sist nevner litteraturen at Big Data gir mulighet for styringsfunksjonen til å jobbe med ustrukturerte datakilder som bilder, lyd og tekstdata (Rikhardson & Yigitbasioglu, 2018). Dette kan for eksempel bety at styringsfunksjonen kan bruke lyddata til overvåking av samtaler med kunder (gitt hensyn til GDPR) eller bilder for å registrere driftsavvik. På denne måten kan styringsfunksjonen oppdage nye trender og sammenhenger i datasettet som den tidligere ikke har kunnet gjøre basert på mindre utdrag av populasjonen.

Dagens forretningsverden krever tilgang til tidsrelevant data, og for å være konkurransedyktig kreves det i større grad at bedrifter klarer å være tilpasningsdyktige (Nielsen, 2018). For å forsyne beslutningstakere med relevant informasjon, stilles det derfor krav til at

styringsfunksjonen må ha fokus på å gjøre prediktive analyser og prognoser som kan hjelpe organisasjoner med å ta gode valg for fremtiden (Nielsen, 2018). Ved å benytte informasjon fra interne og eksterne kilder, kan Big Data hjelpe styringsfunksjonen med å besvare spørsmål knyttet til både hva som har skjedd, hva som vil skje og hva som er den mest optimale løsningen (Appelbaum et al., 2017).

Virksomheter som ønsker å nyttiggjøre seg av Big Data har ofte investert store mengder ressurser for å samle, bearbeide og analysere de store datamengdene, og forventer følgelig avkastning i form av dypere innsikt og bedre kunnskap (Appelbaum et al., 2017). På tross av dette, viser undersøkelser at selv om det er blitt tungt investert i Big Data er det mange selskaper som sliter med å utnytte datamengdene slik at de kan oppnå innsikt som faktisk kan utgjøre en forskjell (Grover et al., 2018). Dette gir et innblikk av at mange organisasjoner finner det utfordrende å jobbe med Big Data og sliter med å hente ut verdier.

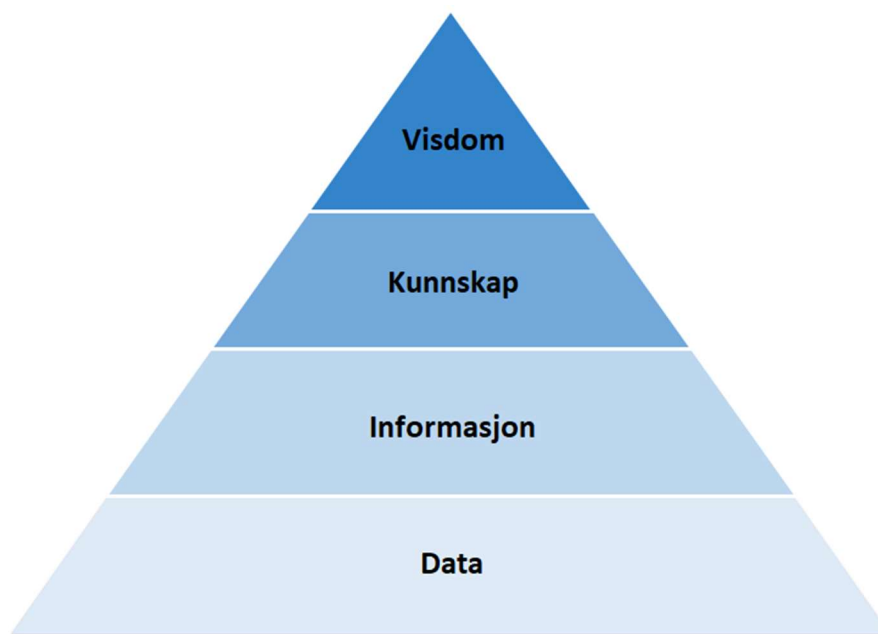
Det kan være mange grunner til at Big Data anses som utfordrende å jobbe med. The Economist (2017) hevder at verdens mest verdifulle ressurs ikke lenger er olje, men data. På en annen side innebærer dette sitatet at data trenger å bli raffinert, strukturert og prosessert før det kan gi verdifulle innsikter (Knudsen, 2020). Dette kan være utfordrende med tanke på kjennetegnene til Big Data; enormt volum, rask omløpshastighet, stor variasjon og usikker kvalitet (Lee, 2017). Med andre ord kreves det stor innsats for å hente ut verdier fra Big Data.

Gitt den store kompleksiteten og ustrukturerte formen av mye av disse dataene kan det virke overveldende å jobbe med disse sammenlignet med tradisjonelle databaser, som gjerne har en mye klarere struktur (Appelbaum et al., 2017). Her kreves det avansert analyse og det å velge ut hvilke data man ønsker å bruke kan være en stor utfordring (Nielsen, 2018).

En videre diskusjon rundt Big Data og styring handler om dataens rolle i beslutningstaking (Nielsen, 2018; Quattrone, 2016; Knudsen, 2020). Ved å benytte Big Data som grunnlag for analyser og beslutninger kan organisasjoner bevege seg mot det som ofte refereres til som 'evidensbasert beslutningstaking' (Holsapple et al., 2014), der data er hovedgrunnlaget for beslutninger fremfor subjektiv erfaring og ekspertise. Utfordringen oppstår imidlertid dersom man har forventning om at data alene gir mening og verdi for virksomheten. Nielsen (2018) hevder her at data og statistiske verktøy i seg selv ikke utgjør noe informasjon, og at det kun er det menneskelige sinn som kan gi mening og relevans til data. Derfor er det viktig at

informasjon blir omgjort til kunnskap før man fatter beslutninger (Nielsen, 2018). Dette støttes av Quattrone (2016) som påpeker at informasjon ikke er det samme som kunnskap. For å oppnå kunnskap må man kombinere informasjon med erfaring, kontekst, tolkning og handling (Nielsen, 2018). Dette sier noe om at gode beslutninger krever kunnskap, altså både tilgang på god informasjon, men også erfaring og riktig tolkning.

En modell som illustrerer forholdet mellom data og kunnskap er «The data-information-knowledge-wisdom hierachry» (DIKW), også kalt visdomspyramiden (Rowley, 2007), som illustrert i figur 2.1. Modellen bygger på den implisitte forutsetningen om at data kan brukes til å skape informasjon. Den antar også at informasjon kan brukes til å skape kunnskap, som videre kan brukes til å skape visdom. Det stilles imidlertid en rekke krav for å komme til dette målet. I det første steget er det viktig at dataene gir mening, og at man har detaljert data som besvarer spørsmål slik som hvem, hva, når og hvor. For å gå videre fra informasjon til kunnskap er man nødt til å relatere dataene til noe, sette de i kontekst eller besvare hvordan noe kan gjøres, såkalt “know-how”. Dette samsvarer godt med hvordan kunnskap er beskrevet i Nielsen (2018), hvor erfaring, kontekst, tolkning og reaksjon er sentralt. Videre resonnerer det godt med annen forskning som påpeker at menneskelig skjønn er viktig også i datadrevne beslutninger (Rikhardson & Yigitbasioglu, 2018). Visdom oppnås først når man klarer å bruke kunnskapen til å ta gode beslutninger (Rowley, 2007). Dette er naturligvis et viktig mål man ønsker å oppnå ved bruk av Big Data i styringsfunksjonen.



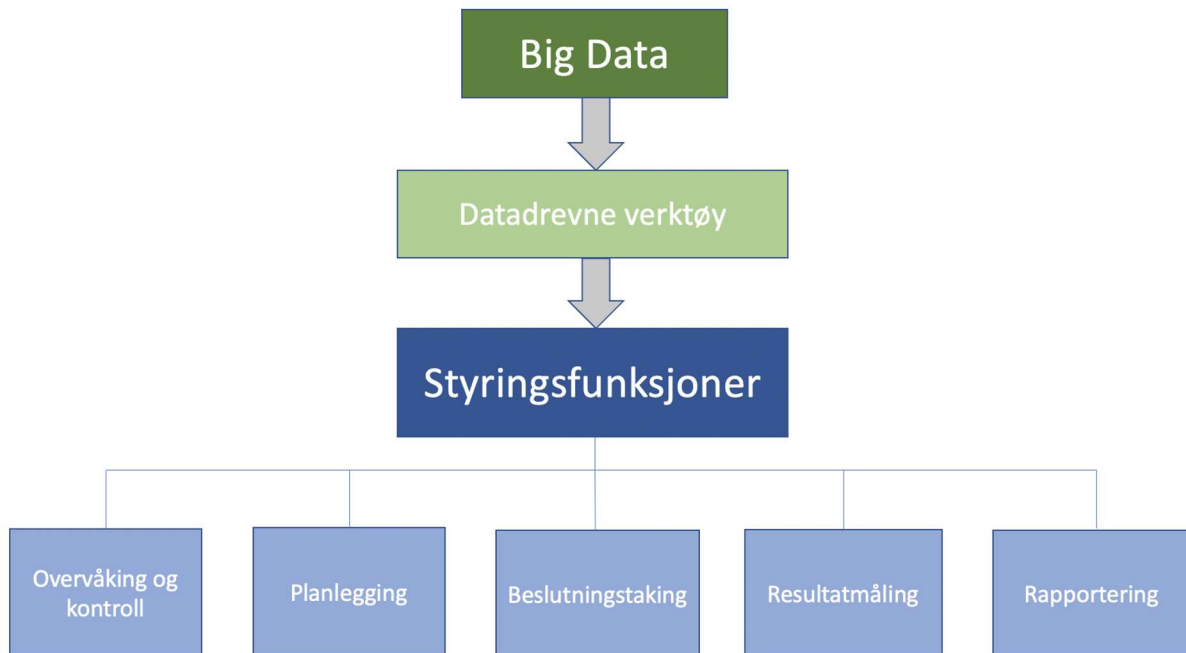
Figur 2.1: Visdomspyramiden (Rowley, 2007)

I forlengelse av diskusjonen knyttet til dataens rolle i beslutningstaking, har flere kommet med advarsler mot datadreven beslutningstaking. Quattrone (2016) advarer mot at et digitalt skift i beslutningstaking vil medføre mindre rom for verdifull diskusjon og menneskets bruk av skjønn. Forfatteren hevder videre at styringsfunksjonen tidligere har bidratt med verdifulle dialoger, og frykter at dette blir nedprioritert og forsvinner helt ettersom mange har forventninger til at Big Data og analytiske verktøy i større grad kan bidra til rasjonell beslutningstaking. Quattrone (2016) hevder derimot at styringsfunksjonen burde sikte på *rimelige* beslutninger og ikke rasjonelle, ettersom fullt ut rasjonelle beslutninger er umulig å få til.

Oppsummert kan vi se at litteraturen viser til at Big Data bidrar med både med store muligheter, men også store utfordringer for organisatorisk styring. Tilgang til store mengder detaljert data gir mulighet til å få et bedre beslutningsgrunnlag, oppdagelse av nye innsikter og dypere forståelse av forretningsprosesser. Samtidig er Big Data vanskelig å jobbe med, spesielt grunnet det enorme volumet og den store variasjonen i datatyper. Den kontinuerlige registreringen av sanntidsdata fra mange ulike kilder medfører at det oppstår spørsmål knyttet til hva man skal gjøre med de enorme mengdene og hvordan man henter ut verdi fra dem. Vi har også sett på forholdet mellom data og kunnskap i beslutningstaking, og noe av kritikken rettet mot datadrevne beslutninger. Generelt virker denne kritikken å handle om at data burde brukes som et informasjonsgrunnlag, hvor menneskets kunnskap og kompetanse fortsatt burde være en sentral del av beslutningstaking.

2.4 Konseptuelt rammeverk

For å oppsummere teoridelen har vi utarbeidet et konseptuelt rammeverk som illustrerer sammenhengen mellom Big Data og styringsfunksjoner. Rammeverket vist i figur 2.2 vil brukes til å strukturere datainnsamlingen og den videre presentasjonen av funn i kapittel 5.



Figur 2.2: Konseptuelt rammeverk for anvendelse av Big Data i styringsfunksjoner

3. Casebeskrivelse

I denne delen skal vi presentere Vann- og avløpsetaten i Bergen kommune som figurerer som case-organisasjon i denne oppgaven. Først presenteres informasjon for å gjøre leseren bedre kjent med Bergen kommune, datasjøen og Vann- og avløpsetaten. Deretter vil vi kort presentere hvilke konkrete styringsfunksjoner som vil være i fokus videre i oppgaven.

3.1 Bergen kommune

Med omkring 286 000 innbyggere er Bergen kommune den nest største kommunen i Norge (SSB, u.d.). Omfanget gjør at kommunen har mye ressurser og stor drivkraft. Dette gir store muligheter for å leve opp til kommunens mål, som er å levere gode og ressurseffektive tjenester til kommunens innbyggere. En stor kommune innebærer også mye kompleksitet, ettersom kundemassen er stor og har mange ulike behov. Innbyggerne består av alle mulige mennesker og er en god representasjon av mangfoldet i Norge. De ansatte i Bergen kommune utgjør i overkant av 15 000 årsverk, enten direkte eller i organisasjoner som er underlagt kommunen (Bergen kommune, u.d.-d). Dette gjør kommunen til en av de største arbeidsgiverne på Vestlandet. Her jobber alt fra sjåførere, lærere, brannmenn, håndverkere og videre, noe som innebærer at det også er et stort spenn i behov fra de ansatte. Kommunen har totalt 22,5 milliarder kroner i årlig driftsbudsjett (Bergen kommune, u.d.-a)

Bergen kommune er en av to norske kommuner som er basert på byparlamentarisme, noe som innebærer at man gjennomfører demokratiske valg for å utforme et bystyre. Bystyret velger et byråd som utgjør det utøvende organet av kommunen. Byrådet er ansvarlig for å utøve bystyrets politiske vedtak. Bystyret og byrådet kan dermed sammenlignes med henholdsvis stortinget og regjeringen. Totalt består byrådet av seks byråder og en byrådsleder, hvor de seks byrådene leder hver sin byrådsavdeling. Disse har ansvar for hver sine etater, virksomheter og resultatenheter, som videre er ansvarlige for den operasjonelle tjenesteleveransen til innbyggerne. Figur 3.1 viser inndelingen av de ulike byrådsavdelingene.



Figur 3.1: Inndeling av byrådsavdelinger (Bergen kommune, u.d.-d)

For å etterleve sitt mål om gode og ressurseffektive tjenester, har kommunen satset tungt på digitalisering, utvikling og bruk av nye teknologiske løsninger for å fornye, forenkle og forbedre sine tjenester (Bergen kommune, u.d.-b). Ifølge byrådet for 2019-2023 skal Bergen kommune være en offensiv foregangskommune for digitale løsninger. Kommunen har derfor vedtatt en digitaliseringsstrategi for perioden 2020-2025 som er inndelt i fire innsatsområder.

For det første skal kommunen være en viktig bidragsyter i arbeidet med å bygge én digital offentlig sektor, hvor blant annet deling og tilgjengeliggjøring av offentlig data er sentralt. Neste innsatsområde er å benytte teknologi og digitalisering til å utvikle en mer bærekraftig tjenesteproduksjon. Dette innebærer blant annet fokus på digital kompetanse og løsninger med gode brukeropplevelser for de ansatte. Det tredje innsatsområde handler virksomhetsutvikling, hvor kommunen skal være en omstillingsdyktig kommune med fremtidsrettede tjenester. Det siste innsatsområdet handler om innovasjon og datadrevet forvaltning, noe som skal oppnås med blant annet trygg deling av data, god datakvalitet og bruk av Big Data.

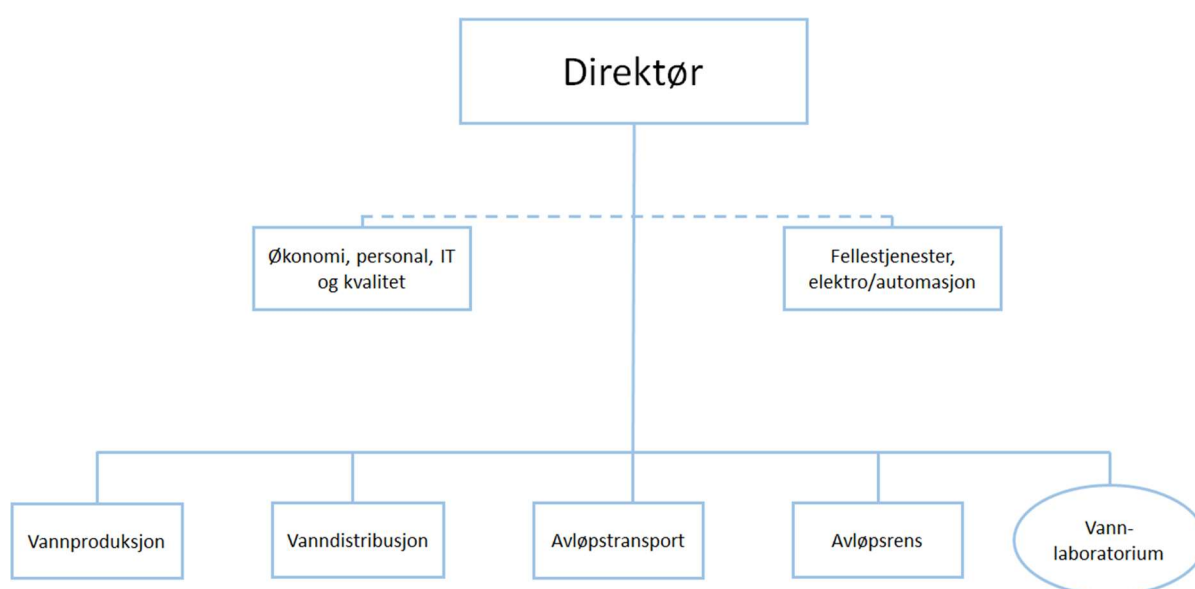
3.2 Vann- og avløpsetaten

Vann- og avløpsetaten (VA-etaten) er organisert under Byrådsavdelingen for klima, miljø og byutvikling. Med sine omtrent 95 ansatte sørger VA-etaten for at kommunens ansvar i henhold til vannforsyning og avløpshåndtering ovenfor abonnenter og myndigheter blir ivarettatt (Bergen kommune, u.d.-c). Etaten har to hovedmål; For det første skal de dekke etterspørselen etter vann med drikkevannskvalitet, og for det andre skal de ta hånd om avløpsvann slik at miljøskade og sjenerende forhold ikke oppstår (Bergen kommune, 2020). Denne virksomheten

er organisert etter bestiller-/utfører modellen hvor drift og vedlikehold utføres av Bergen Vann KF (Bergen Vann), på oppdrag fra VA-etaten.

Totalt forvalter VA-etaten en infrastruktur bestående av om lag 1000 km vannledninger, 1200 km avløpsledninger, 5 avløpsrenseanlegg, 40 høydebasseng og 200 pumpestasjoner. Ettersom store deler av disse komponentene er under bakken, er driften av nettet nokså komplisert. Etaten har flere definerte mål, blant annet å ha fokus på innovasjon og smarte løsninger innen sektoren. Videre har etaten flere operasjonelle mål, som for eksempel å redusere lekkasjetapet fra dagens omtrent 31 % til under 20 % frem til 2028, samt å fornye 0,7 % av vannledningsnettet per år.

Bergen Vann KF har ansvar for drift og vedlikehold av nettet i Bergen kommune, og er et selvstendig kommunalt foretak med eget styre som eies 100 % av kommunen. Denne organisasjonen har rundt 150 ansatte (Bergen kommune, u.d.-c), noe som tilsier at begge organisasjonene samlet har omtrent 245 ansatte. Bergen Vann er organisert i fire ulike seksjoner; vannproduksjon, vanddistribusjon, avløpstransport og avløpsrens. Blant driftsseksjonen er det også vannlaboratorium som gjør analyser og test av blant annet vannkvaliteten. I tillegg til driftsseksjonen har organisasjonen støttefunksjoner som økonomi, personal, IT og fellestjenester. Et organisasjonskart av Bergen Vann ser vi i figur 3.2.



Figur 3.2: Bergen Vann organisasjonskart (Bergen Vann, u.d.)

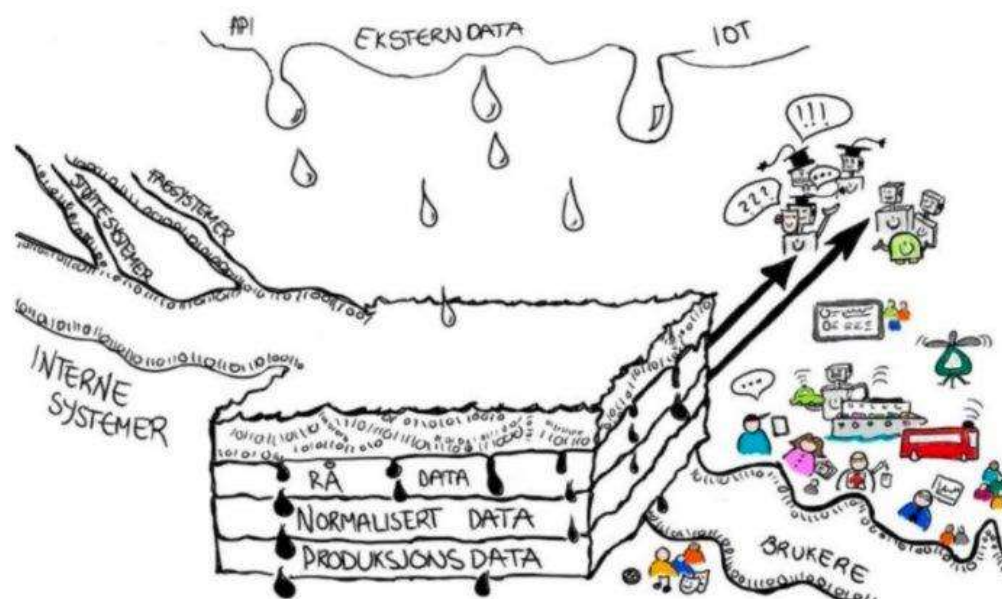
Bergen Vann ble skilt ut av VA-etaten i 2004. Bakgrunnen for dette var at VA-etaten skulle ha et mer langsiktig fokus, mens Bergen Vann skulle ta seg av de operasjonelle oppgavene knyttet til drift og vedlikehold (Bergen kommune, u.d.-c). Det var også et politisk ønske om at denne driften skulle konkurranseutsettes, og Bergen Vann har derav levert VA-tjenester også for private oppdragsgivere. Det er derimot fattet vedtak i bystyret om at disse to organisasjonene skal slås sammen fra høsten 2021 under navnet 'Etat for vann- og avløpstjenester'. Med tanke på at VA-etaten og Bergen Vann KF har et tett samarbeid, og det faktum at de nå skal slås sammen igjen fra høsten 2021, har vi for enkelhetens skyld valgt å slå sammen disse to og referere til dem som Vann- og avløpsetaten (VA-etaten) videre i oppgaven.

3.3 Datasjøen

I 2019 lanserte Bergen kommune datasjøen *Lungegårdsvannet* (Bouvet, 2019). Datasjøen driftes av Seksjon Digitalisering og Innovasjon (SDI), underlagt Byrådsavdeling for finans, næring og eiendom. Hensikten med datasjøen er å forenkle tilgjengelighet og deling av data slik at den kan brukes bedre av aktører både internt og eksternt (Bergen kommune, u.d.-b). Kommunen beskriver datasjøen på følgende måte:

«Datasjøen er en metode for lagring av alle former for data og kan sammenlignes med et sentralt datalager for alle typer data: strukturerte og ustrukturerte, både dokumenter og logger, bilder, lyd og video. Den kan være en kilde til alle data innenfor et område med mulighet for tilgang for flere, og et verktøy for effektivisering; læring, planlegging, utforskning av muligheter og et viktig grunnlag for maskinlæring og kunstig intelligens. En datasjø kan legge til rette for effektiv og standardisert datadeling, med sikre tilgangsmekanismer.»
(Bergen kommune, u.d.-b)

Beskrivelsen av Bergen kommune sin datasjø virker å passe godt med definisjoner av Big Data som vi har sett på tidligere under teorien. Datasjøen virker å være et klart og lokalt uttrykk for konseptet Big Data som en kilde og lagringsplass for alle former for data, både strukturerte og ustrukturerte. Figur 3.3 viser en illustrasjon av konseptet med datasjøen.



Figur 3.3: Illustrasjon av datasjøen til Bergen kommune (Illustrert av Kjersti Haukeland Eiken, SDI, Bergen kommune).

Bergen kommune virker å ha store ambisjoner knyttet til dette prosjektet og ønsker at datasjøen vil støtte opp under målet om å gjøre regionen mer datadrevet. Formålet med datasjøen er tredelt. For det første skal datasjøen tilby åpne offentlige data, både for kommunens egne enheter og for private aktører i kommunen. For det andre skal datasjøen tilrettelegge og være en driver for innovasjon på tvers av byrådsavdelinger slik at man får forbedring og utvikling av kommunens tjenester. For det tredje skal datasjøen tilrettelegge for bedre innsikt og beslutningsstøtte gjennom bruk av data, noe som igjen kan muliggjøre en mer effektiv drift av kommunens etater. Dette viser også at Big Data kan brukes til styringsformål.

Datasjøen kan videre bli et viktig element for å følge opp Regjeringens strategi for kunstig intelligens, som ble fremlagt i 2020. I et høringsutkast påpeker kommunen at det ligger et stort potensial for å utnytte store datamengder i datasjøen til å finne mønstre og sammenhenger gjennom maskinlæring, som kan gi nyttig informasjon- og beslutningsstøtte. Det er laget noen styringsprinsipper for datasjøen som bestemmer hvordan den skal forvaltes. Blant annet leveres tjenestene fra datasjøen i dag gjennom såkalte *brukerhistorier*, som innebærer at tjenestene leveres på behovsbaserte bestillinger fra brukerne. Dette skal bidra til at behovene til brukeren står i sentrum og at tjenestene som leveres har verdi. Dette kan være en ny enkeltstående tjeneste, en del av en ny tjeneste eller som en berikelse av en eksisterende tjeneste. Slike brukerhistorier kan bestilles av etater og andre aktører, og blir produsert av kommunens eget datasjøteam i henhold til egne retningslinjer.

VA-etaten var tidlig ute som brukere av tjenester fra datasjøen og bidrar blant annet med data fra hele 13 000 sensorer fra etatens styringssystem SCADA (Mekki, 2020). Her registreres og lagres sensordata hvert minutt med ulike måleverdier slik som vannmengde, trykk og temperatur. Fra de første linjene med data ble plassert i august 2018, har VA-etaten årlig lagret over 7 milliarder måleverdier i datasjøen (Mekki, 2020). Dette gir et bilde av datavolumet som VA-etaten har i datasjøen.

Grunnlaget for VA-etatens motivasjon til å bli med i datasjøen var tredelt (Mekki, 2020). For det første ønsket de å lagre og ta vare på detaljerte måledata over tid slik at det kan brukes til analyse og maskinlæring. For det andre ønsket de å gi tilgang til måledata for eksterne aktører, uten at de gikk på bekostning av sikkerheten i styringssystemet. Den tredje motivasjonen var å tilgjengeliggjøre måledata for driftsoperatører og andre feltarbeidere gjennom et nettbrett-basert feltverktøy. Dette gir tilgang til måledata i sanntid for arbeidere ute i feltet. Ifølge IT-sjefen i VA-etaten, har alle disse tre målsetningene blitt oppfylt (Bouvet, 2019). Videre har etaten ambisjoner om å utvikle flere løsninger ved hjelp av datasjøen, blant annet ved å trekke inn data fra flere kilder som for eksempel værdata og havnivå-data.

3.4 Styring i Vann- og avløpsetaten

I teorigapittelet om styringsfunksjoner 2.2.1 kom det frem at styringsfunksjonene kan variere med hensyn til kontekst og hvilken situasjon organisasjonen befinner seg i. Det vil derfor være hensiktsmessig å ta en vurdering av disse funksjonene i lys av case-organisasjonen når vi bestemmer hvilke funksjoner som skal fokuseres på i denne studien.

Fra kapittel 2.2.1 kom det frem at funksjoner som planlegging, overvåking, kontroll, resultatmåling, kostnadsføring, rapportering, transaksjonsanalyse og beslutningstaking ble fremhevet som sentrale funksjoner. Sett opp mot caset i denne oppgaven, som er VA-etaten i Bergen kommune, virker mange av disse styringsfunksjonene å være relevante. Fra innledende samtaler med VA-etaten oppfattet vi blant annet at oppgaver som overvåking og kontroll av vann- og avløpsnett, planlegging av vedlikehold og fornying av rør, samt beslutningstaking knyttet til hvilke driftstiltak som skal iverksettes, var sentrale styringsfunksjoner for etaten. I tillegg til disse var intern og ekstern rapportering til blant annet myndigheter, samt resultatmåling av diverse KPI'er, noe som fremstod som sentrale funksjoner.

Videre fikk vi et inntrykk av at beslutningen som ble tatt i VA-etaten hadde stor bredde, men at det i hovedsak handlet om operasjonelle beslutninger og ikke så mye finansielle beslutninger. Dette kan ha sammenheng med at organisasjonen er en offentlig virksomhet, hvor de finansielle styringsoppgavene kanskje ligger hos andre enheter. VA-etaten er som nevnt ikke en profittmaksimerende organisasjon, men en driftsorganisasjon som skal ivareta stabile og gode tjenester for abonnentene sine. På bakgrunn av dette er det forståelig at etaten har et stort fokus på operasjonelle prosesser. Vårt fokus i denne oppgaven ligger derfor på de operasjonelle styringsfunksjonene, og ikke de finansielle.

Etttersom fokuset i denne oppgaven er på de operasjonelle funksjonene, vil styringsfunksjoner her konseptualiseres som overvåking og kontroll, planlegging, beslutningstaking, resultatmåling og rapportering. Styringsfunksjoner i litteraturen som er av finansiell karakter, slik som kostnadsføring og transaksjonsanalyse, vil dermed ikke inkluderes blant de sentrale styringsfunksjonene i denne oppgaven.

4. Metode

I dette kapittelet vil vi redegjøre for de metodiske valgene som er gjort for å besvare oppgavens problemstilling. Først vil valg av forskningstilnærming, forskningsdesign, datainnsamling og dataanalyse beskrives, før diskusjonen går videre på betraktninger rundt kvaliteten på den metodiske tilnærmingen. Kapittelet avsluttes med vurderinger rundt etiske utfordringer og metodiske begrensinger av oppgaven.

For å belyse forskningsspørsmålet *Hvordan anvendes Big data i organisatorisk styring* har vi valgt å basere oss på en kvalitativ forskningsmetode. Ifølge Saunders, Lewis & Thornhill (2016), brukes begrepet «kvalitativt» ofte som et synonym for enhver teknikk for dataanalyse eller prosedyre for analyse av data som genererer eller bruker ikke-numerisk data. Ved en slik metode kreves det at data samles inn med et åpent sinn hvor man er mottakelige for nye og uforutsette funn (Qu & Dumay, 2011; Scapens, 2004). Vi anser derfor dette som en hensiktsmessig metode for vår oppgave, da fenomenet vi undersøker er relativt bredt og komplekst.

4.1 Valg av metode

4.1.1 Forskningstilnærming

Oppgaven har en abduktiv forskningstilnærming, som er en kombinasjon av deduktive og induktive forskningstilnærminger (Saunders et al. 2016). I en slik tilnærming anerkjennes det at man er påvirket av noen teoretiske konsepter før man går ut i felt (Timmermans & Tavory, 2012). Dette kan anses å være en passende tilnærming for denne utredningen ettersom fenomenet organisatorisk styring er relativt bredt og komplekst. Ved å sette noen løse teoretiske rammer for fenomenet, gir vi noen strukturer til datainnsamlingen som kan være nyttige. Selv om vi gjorde oss kjent med enkelte teoretiske konsepter om fenomenet på forhånd, holdt vi løst fast på disse og var åpne for å gi slipp på dem dersom funn indikerte at de ikke var relevante. På denne måten lar tilnærmingen oss gå frem og tilbake mellom teori og empiri, og kan lede til ny teori eller modifikasjon av eksisterende teori (Saunders et al., 2016).

4.1.2 Forskningsdesign

For å besvare oppgavens problemstilling har vi valgt å basere oppgaven på et eksplorerende forskningsdesign. I et slikt forskningsdesign stiller man åpne spørsmål for å oppdage hva som skjer og få innsikt i et tema av interesse (Saunders et al., 2016, s. 174). Big data er et fenomen som er relativt nytt og er derav nokså lite utforsket generelt sett. Anvendelse av Big data i styringsfunksjonen er langt mindre utforsket, særlig som empiriske studier (Rikhardson & Yigitbasioglu, 2018). Etersom dette er et relativt nytt, ustrukturert og komplekst fenomen, virker en eksplorativ studie å være passende for å kunne videreutvikle et teoretisk perspektiv om fenomenet.

4.1.3 Forskningsstrategi

Med forskningsstrategi menes en generell plan for hvordan forskningsspørsmålet skal besvares, og er ifølge Saunders et al. (2016) en metodisk link mellom forskningsfilosofien og de påfølgende valgene av forskningsmetodene som brukes for å samle inn og analysere data. For å få et godt innblikk i hvordan data brukes i styringsfunksjonen ser vi det som hensiktsmessig å se på fenomenet i lys av en enkelt casestudie. Casestudier benyttes gjerne på et mindre utvalg hvor målet er å søke etter dybdekunnskap om et tema eller fenomen (Saunders et al., 2016). Alternativt kunne vi gjennomført et multi-casestudie av flere virksomheter for å sammenligne fenomenet på tvers. Men ettersom vi i forskningsspørsmålet ønsker å belyse *hvordan* Big Data anvendes i styring, virker en casestudie å være godt egnet ettersom vi får undersøkt fenomenet i dybden i en bestemt kontekst. Dette kan bidra til å få frem detaljer og nyanser som gjerne kan forsvinne dersom man gjør en multi-casestudie med mer aggregerte data. På grunn av begrenset tid og ressurser, som ofte er tilfellet for masteroppgaver, gjennomførte vi en tverrsnittstudie. Dette vil si at vi undersøker fenomener, mønstre og sammenhenger på et bestemt tidspunkt (Grønmo, 2021).

4.2 Datainnsamling

Datamaterialet for studien er basert på både primær- og sekundærdata. Saunders et al. (2016) peker på sekundærdata som en rimelig og ikke-påtrengende måte å samle inn data på, men påpeker at den ofte er produsert med et formål som ikke nødvendigvis er relevant for

forskningsspørsmålet man undersøker. Allikevel mener vi at sekundærdata, som for eksempel interne instruksjoner og offentlige presentasjoner av digitale løsninger, kan være nyttig for å skaffe mer innsikt i temaet vi undersøker.

Primærdataen i studien er samlet inn gjennom å gjennomføre 8 semi-strukturerte dybdeintervju, i stor grad basert på teori rundt MA-funksjonen. Med unntak av ett intervju, var alle intervjuene individuelle. Dette unntaket ble gjort på bakgrunn av ønske fra to av informantene. Vi ser individuelle dybdeintervju som hensiktsmessig da det kan tenkes at informantene har ulike erfaringer om hvor godt de ulike løsningene fungerer og det kan være stor variasjon i bruk. En fordel med en slik intervjuform er også at spørsmålene og rekkefølgen kan tilpasses etter behov (Saunders et al., 2016).

4.2.1 Dokumentanalyse

I tillegg til innsamlingen av primærdata ble det også samlet inn sekundærdata gjennom søk etter informasjon på kommunens nettsider, samt undersøkning av dokumenter utdelt fra våre kontakter i kommunen. Flere av disse dokumentene lå også åpent ute på nett og var i liten grad konfidensielle. Dokumentene bidro til å få en rask og god innsikt i aktuelle digitaliseringsinitiativ knyttet til datasjøen og VA-etaten. Dette var nyttig ettersom vi visste lite om disse initiativene før starten av prosjektet.

4.2.2 Utvalg

Utvalget i studien er basert på et bekvemmelighetsutvalg. Gjennom vår veileder kom vi i kontakt med Bergen kommune og personer i Seksjon Digitalisering og Innovasjon (SDI), som videre satt oss i kontakt med personer i VA-etaten. I innledende samtaler ble det diskutert forutsetninger og formål med oppgaven, før kontaktpersonen i VA-etaten foreslo 10 aktuelle informanter som kunne bidra med innsikt i problemstillingen. Av de aktuelle informantene som ble foreslått, ville 8 av 10 bidra med intervju. Samtidig ønsket én av informantene å ha med seg en kollega, noe som førte til at det var to informanter til stede på første intervju. Vi endte derfor opp med totalt 9 informanter til denne studien. De hadde varierte stillinger, og bestod av blant annet seksjonsledere, prosjektledere og optimaliserings-ingeniører. I tabell 4.1 er det en oversikt over informantene, hvilken seksjon de tilhørte, og lengden på intervjuene.

| Informanter | Seksjon | Intervjuets lengde |
|--------------------|----------------------------------|---------------------------|
| Informant 1a | Avløpstransport/Planavdeling | 51 minutt |
| Informant 1b | Planavdeling | 51 minutt |
| Informant 2 | Vanndistribusjon | 52 minutt |
| Informant 3 | Avløpsrens | 43 minutt |
| Informant 4 | Vanndistribusjon | 50 minutt |
| Informant 5 | Vanndistribusjon | 48 minutt |
| Informant 6 | Avløpstransport | 50 minutt |
| Informant 7 | Avløpstransport | 43 minutt |
| Informant 8 | Vanndistribusjon/Avløpstransport | 38 minutt |

Tabell 4.1: Informanter

4.2.3 Intervjuprosessen

Grunnet situasjonen med koronapandemien utførte vi alle intervjuene digitalt, ved hjelp av tjenesten Microsoft Teams. Informantene fikk på forhånd tilsendt et informasjonsskriv og samtykkeskjema som blant annet inneholdt informasjon om oppgavens formål og deres rettigheter i forhold til persondata som ble samlet inn. Både i samtykkeskjemaet og i samtale før intervjuet ble det innhentet samtykke til å benytte lydopptak av intervjuene. Informasjonsskrivet opplyste om at intervjuene ville vare i rundt 30-45 minutter, mens noen av intervjuene varte noe lenger ettersom informantene hadde mye de ville dele. Før intervjuet startet vi med en kort presentasjon av oss selv og vår bakgrunn, samt informasjon om personvern og hvordan datamaterialet skulle brukes. Under intervjuene hadde én av oss ansvar for å lede intervjuet mens den andre hadde hovedansvar for å ta korte notater, ordne med det tekniske rundt lydopptaket og stille eventuelle oppfølgingsspørsmål. Etter intervjuene ble datamaterialet transkribert og anonymisert i henhold til NSDs retningslinjer.

4.2.4 Dataanalyse

Intervjuene ble løpende transkribert gjennom datainnsamlingsperioden for å sikre så presis gjengivelse av meningsinnholdet i informantenes utsagn som mulig. Dette var også hjelpsomt for å få et overblikk over datamaterialet som vi videre kunne bruke i arbeidet med koding og kategorisering. Det ble gjennomført en tematisk analyse, som anses som den mest grunnleggende metoden for å analysere kvalitative data. Dette er en fleksibel metode for analyse av datasett av ulike størrelse og kan ifølge Saunders et al. (2016) lede til rike beskrivelser, forklaringer og teoretisering. Selve gjennomføringen av analysen begynte med å gjennomgå materialet flere ganger, og identifisere enkelte meningsinnhold, eller *koder*. Dette ble gjort hver for oss for å unngå å påvirke hverandre. Etter kodingen startet prosessen med å lete etter *temaer* og sammenhenger. Temaer er en bred kategori som inneholder relaterte koder og indikerer idéer som kan være nyttig for å besvare forskningsspørsmålet (Saunders et al., 2016). Dette ble videre raffinert og gav oss et godt grunnlag for å beskrive våre funn.

4.3 Evaluering av utredningen

I kvantitative studier vurderes gjerne kvaliteten av forskningen basert på reliabilitet og validitet. Flere forskere har imidlertid påpekt at en slik vurdering er vanskelig innenfor eksplorative casestudier (Berry & Otley, 2004; Scapens, 2004). For å vurdere kvaliteten av vår studie velger vi dermed å basere diskusjonen på evalueringskriteriene som beskrevet i Scapens (2004), samt trekke inn vurderinger fra det mer kjente rammeverket fra Lincoln og Guba (1989).

4.3.1 Prosessuell reliabilitet

For å vurdere forskningskvalitet i casestudier er det viktig å vite at forskerne har benyttet passende og pålitelige forskningsmetoder og prosedyrer, noe Scapens (2004) definerer som *prosessuell reliabilitet*. For å best mulig støtte oppunder dette beskrives våre forskningsmetoder og prosedyrer i detalj i dette kapitlet, og gjennom hele studien har det vært viktig å gi leseren et godt revisjonsspor for å se hvordan dataen er samlet inn og analysert i studien. Dette innebærer utdypende beskrivelser hvor det er forklart nøye og ærlig hvordan man har gått frem og hvilke funn man har gjort, noe som gjerne betraktes som «*thick*

description» (Lincoln & Guba, 1985). Vi gjorde oss også godt kjent med VA-bransjen og med MA-litteraturen i forkant av gjennomføringen av dybdeintervjuene. Dette gjorde også at intervjuguiden var godt gjennomtenkt på forhånd.

En av fallgruvene innenfor dybdeintervju er at informantene ikke nødvendigvis svarer ærlig på alle spørsmål eller ikke besitter riktig kunnskap (Qu & Dumay, 2011). Vi startet intervjuene med å informere om formålet med prosjektet og stilte deretter noen enkle innledende spørsmål. Dette steget kan ifølge flere forfattere være viktig for å skape tillit og kan dermed stimulere til at informantene snakker mer åpent (Qu & Dumay, 2011). I tillegg var vi bevisst på risikoen om at vi som forskere kan påvirke informantene dersom vi ikke er påpasselige på egen opptreden og objektivitet (Qu & Dumay, 2011). Vi var derfor opptatt av å gi informantene god tid mens de snakket, slik at de ikke ble avbrutt eller at vi selv fullførte setningene deres. Vi brukte også egne webkamera i alle intervjuene, og var opptatt av at vårt eget kroppsspråk ikke skulle påvirke deltakerne, for eksempel i situasjoner der svarene informantene gav ikke sammenfalt med våre forventninger. For å sikre på at vi kunne korrekt gjengi det som kom fram i intervjuene benyttet vi lydopptak som vi videre transkriberte.

4.3.2 Overførbarhet

Scapens (2004) påpeker at statistiske generaliseringer er åpenbart problematiske i en casestudie setting. I slike studier er dermed mange forskere som ser bort i fra hele generaliseringsbegrepet, og fokuserer heller på *overførbarhet* av funnene fra en kontekst til en annen og *fittingness* angående graden av sammenlignbarhet av forskjellige kontekster (Scapens, 2004). På bakgrunn av vårt valg av forskningsdesign, samt at vi studerer et fenomen som er situasjonsspesifikk, vil det være vanskelig å oppnå direkte overførbarhet av dette caset. Vi ser likevel ikke bort fra at funnene våre kan brukes som basis for videre forskning, og dermed bidra til mer generaliserbar teori i fremtiden. Under dette punktet har det også vært viktig med transparens og en såkalt «thick description», slik at det blir lettere for andre å vurdere om funnene våre kan være anvendes i en annen aktuell kontekst.

4.3.3 Kontekstuell validitet

I stedet for *intern validitet* er det blitt argumentert for at casestudier bør vurderes i henhold til *kontekstuell validitet*. Scapens (2004) beskriver dette som en vurdering i forhold til

kredibiliteten til funnene og de resulterende konklusjonene. I vår oppgave har datatriangulering vært et viktig virkemiddel for å styrke dette. Vi har sammenlignet deler av dataen vi innhentet gjennom dybdeintervjuene med offentlig informasjon på internett slik som Powerpoint-presentasjoner, organisering og retningslinjer i forhold til IT-prosjekter. Det ble også gjennomført flere ustrukturerte diskusjoner om prosjektet med ulike personer i Bergen kommune, både før og etter innsamling av primærdata. I og med at intervjuguiden i stor grad var konsentrert rundt den enkeltes arbeidsoppgaver og hvordan den enkelte bruker data i sitt arbeid, blir denne trianguleringen i noe mindre betydning. Den har likevel vært nyttig for å få et helhetlig innblikk i hvordan VA-etaten bruker data i sine styringsfunksjoner. I analysen var vi opptatt av at funnene reflekterte informantenes tanker og ikke ble farget av våre egne meninger. Diskusjonen rundt funnene ble gjort i fellesskap for å unngå personlig forutinntatthet eller feiltolkninger.

4.3.4 Selektiv plausibilitet

En annen fallgrube innenfor kvalitative studier er å være selektiv i hvilket datamateriale man inkluderer i analysen. Scapens (2004) hevder at slik *selektiv plausibilitet* skjer når data som passer til forskerens teori tas med, mens det som ikke passer blir ignorert. Det at fenomenet vi undersøker er lite utforsket tillot oss å være åpne for ulike oppfatninger, noe som gjorde oss mindre tilbøyelige til å foretrekke visse synspunkter. Vi har allikevel vært oppmerksomme på denne mulige faren i presentasjonen av analysen og våre funn. Denne faren har også blitt redusert gjennom valg av forskningsdesign, da vi gjennom studien har vært opptatt av å finne varierte oppfatninger.

4.4 Ethiske betraktninger

Ettersom vår metode for datainnsamling i stor grad har vært basert på dybdeintervjuer, har vi måttet gjøre flere etiske vurderinger knyttet til personvern og databehandling. Det er viktig at forskningsdesignet ikke utsetter informantene for forlegenhet, smerte, skade eller andre ulemper (Saunders et al., 2016). Dette har vi tatt under betraktning gjennom hele studien, og det har vært viktig å sikre informantene adgang til full anonymitet og innsikt i data. Før datainnsamlingen sendte vi prosjektet inn til vurdering av Norsk Senter for Forskningsdata (NSD). Prosjektet, med tilhørende samtykkeskjema, ble godkjent før vi begynte med intervjuene. Vi har tatt hensyn til regelverk angående databehandling, blant annet angående datasikkerhet.

4.5 Metodiske begrensninger ved studien

Som presisert i vurderingen av evalueringen av utredningen finnes det enkelte metodiske svakheter og begrensninger som kan ha påvirket studien. I et casestudie er man alltid begrenset til én enkelt setting, noe som kan medføre at funn varierer fra andre settinger. Dette kan følgelig redusere muligheten for at våre konklusjoner er gyldige i andre situasjoner eller utvalg. Vi vil imidlertid ikke utelukke at funnene som er gjort kan ha likhetstrekk i VA-etater i andre norske kommuner eller andre lignende organisasjoner. I tillegg kan det påpekes at casen med fordel kunne blitt utført som en longitudinell studie fremfor en tverrsnittstudie, da fenomenet er i konstant og rask utvikling

5. Analyse og funn

I denne delen av oppgaven skal vi presentere funn og analysere disse. Funnene er hentet ut primært basert på intervjuer med informantene, men også fra innledende møter med VA-etaten og datasjøen, samt noe sekundærdata fra diverse interne rapporter og dokumenter som vi har fått tilgang til. For å strukturere datainnsamlingen har vi, som vist under kapittel 2.4, utarbeidet et konseptuelt rammeverk som illustrerer anvendelsen av Big Data i styring. Styringsfunksjoner har vi som beskrevet i kapittel 3.4, valgt å dele opp i funksjonene overvåking og kontroll, planlegging, beslutningstaking, resultatmåling og rapportering. Strukturen i analysen følger derfor en gjennomgang av funn knyttet til hvordan Big Data anvendes i de nevnte styringsfunksjonene.

5.1 Anvendelse av Big Data til overvåking og kontroll

Overvåking og kontroll i denne oppgaven anser vi å handle om overvåking av operasjonelle driftsprosesser og identifisering av driftsavvik i disse prosessene. For å drive overvåking og kontroll over operasjonelle prosesser, er tilgang på god informasjon avgjørende. I de neste avsnittene skal vi se nærmere på hvordan Big Data anvendes i denne styringsfunksjonen.

Funn: Sensordata i sanntid anses som ‘synet’ til VA-etaten

I VA-etaten benyttes det et styrings- og overvåkingssystem som heter Cactus. Dette er et såkalt SCADA-system, en forkortelse for *Supervisery Control And Data Aquisition*, og brukes for å overvåke og kontrollere diverse operasjonelle prosesser. Tilkoblet VA-etaten sitt SCADA-system er mer enn 13 000 sensorer som fanger opp og registrerer sanntidsdata hvert minutt fra ulike komponenter i ledningsnettet, slik som avløpspumper, vann- og avløpsledninger, og høydebasseng. Sensorene sammenfatter flere forskjellige typer måleparametere slik som mengde, nivå, trykk, og temperatur. Generelt virker informantene å være klar på at tilgang til sanntidsdata fra dette systemet har stor verdi for overvåking og kontroll av ledningsnettet. Informant 6 forteller hvor viktig dette systemet er for å holde oversikt over prosesser, og at det anses som ‘synet’ til VA-etaten:

«Cactus er synet vårt utad. Siden alt vi gjør er under bakken så er det vanskelig å ha kontroll uten å bruke teknologi og sensorer som kan fortelle oss hva som skjer. Cactus er vårt SCADA system som viser kurver, bilder og informasjon som vi er avhengig av. Her følger vi den daglige oversikten over at alt er som det skal.»

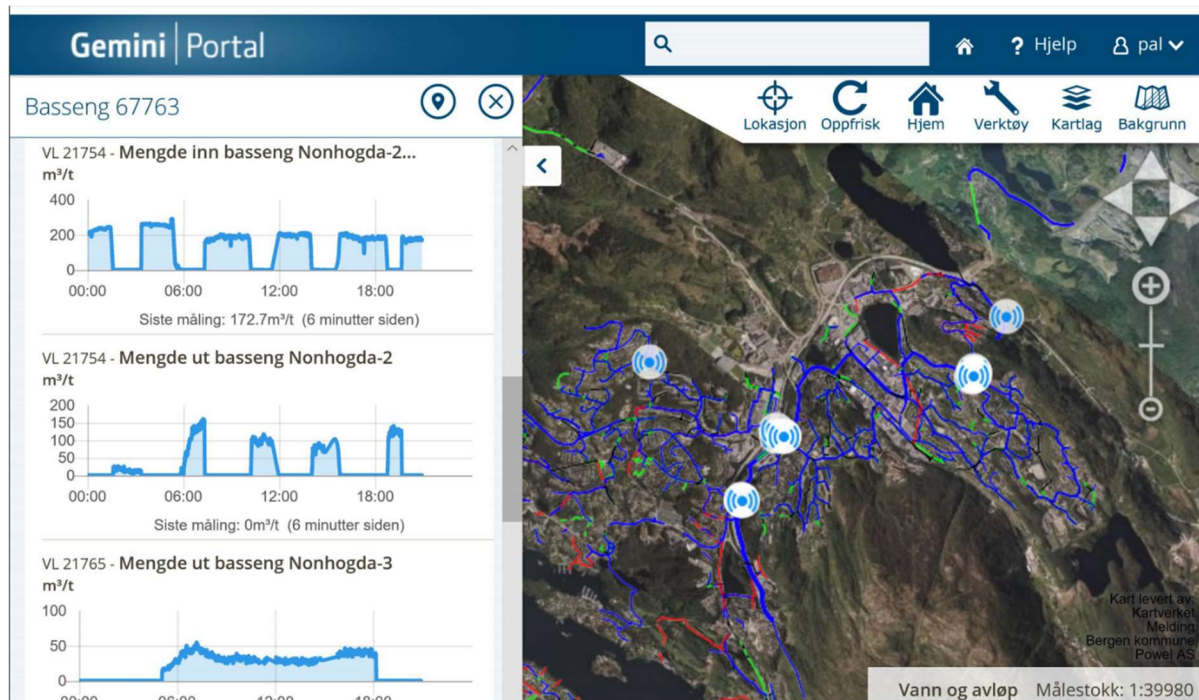
Informant 2 støtter dette og påpeker videre at tilgang til sanntidsdata fra sensorer gir god monitorering av nettet:

«Vi har rundt 1000 km med vannledning, 70 pumpestasjoner og 40 høydebasseng, og alle har relativt god monitorering gjennom sensorer som gir oss data inn i SCADA-systemet som vi bruker for å overvåke nettet. Det hjelper oss til å se at vannleveransen går som den skal til enhver tid. Sånn sett kan du si at de dataene som flyter inn i systemet brukes for å kvalitetssikre driften, og rett og slett sluttleveransen til kunden. (...) Det er klart at uten data så hadde vi vært helt blinde.»

Fra sitatene ser vi at sensordata fra overvåkingssystemet ansees å være 'synet' til VA-etaten. Uten dette virker de å være 'blinde' til hva som skjer rundt omkring på nettet, og avhengigheten til teknologi er tilsynelatende stor. Konkret brukes dette til å 'kvalitetssikre driften', altså en tydelig kobling til styring.

Funn: Big Data anvendes til å overvåke ledningsnettets i sanntid gjennom visualiseringsløsninger

Tidligere litteratur har hevdet at visualiseringsverktøy er et viktig hjelpemiddel for styring. Grunnen til dette er at personer som driver med styring har behov for å få fremstilt informasjon på en rask og effektiv måte, og gode visualiseringer av data kan være effektive til dette formålet. VA-etaten i Bergen kommune har i løpet av de siste par årene fått sammenkoblet sanntidsdata fra SCADA-systemet med et kartsystem utviklet av leverandøren Powel. Kartverktøyet heter *Gemini Portal* og viser oversikt over ulike komponenter i ledningsnettets plassert på et kart, samt måleverdier i sanntid fra de ulike komponentene. En illustrasjon av dette verktøyet ser vi i figur 5.1.



Figur 5.1: Kartverktøyet Gemini Portal (Powel)

Informantene virker å være begeistret over muligheten dette kartverktøyet gir for overvåking og kontroll. Bruk av kartverktøy er ikke noe nytt, men tilgjengeliggjøring av sanntidsdata fra sensorer virker å gi nye dimensjoner til visualiseringen. Istedenfor å kun se hvor de ulike komponentene er lokalisert på kartet, kan man nå hente ut måleverdier i sanntid. Dette kan bidra til å gi et bedre bilde av hva som skjer under bakken til enhver tid. I tillegg til sanntidsdata, viser kartet informasjon om de ulike komponentene slik som alder, størrelse, dimensjoner og driftsmeldinger. Informant 7 beskriver dette verktøyet som en milepæl i databruk hos VA-etaten:

«Den største forbedringen har vært å få fremstilt sanntidsdata i Gemini Portal. Da får man dataen ut til driftsoperatørene, og til alle som trenger tilgang egentlig. Det tror jeg er en milepæl i bruken av data hos VA-etaten. Så det at man på et webgrensesnitt kan hente ut sanntidsdata om pumper og ledninger, er veldig verdifullt, spesielt for feltarbeiderne og driftsoperatørene ute. (...) Det har gitt oss en helt annen kontroll enn tidligere.»

Informanten fremlegger her også et sentralt poeng, som for øvrig støttes av flere andre informanter, nemlig at data gjennom dette verktøyet nå er tilgjengeliggjort for driftsoperatører ute i felten. Dette kan bidra til mer effektiv informasjonsdeling, noe informant 4 belyser ved å fortelle hvordan dette ble gjort før:

«Det at vi har fått data tilgjengeliggjort ute i felten har gjort at vi ikke er avhengig av folk på kontoret. Før måtte vi ringe inn til kontoret for å spørre om de kunne lese av en måler eller se på forbruket. Da var de på kontoret opptatt til kun å sitte og lese gjennom disse dataene.»

Tidligere måtte altså driftsoperatørene ute i felten ringe inn til kontoret for å be dem lese opp forskjellige typer måledata som de trengte. Dette er ressurskrevende både med hensyn til tid, men også i forhold til antall personer som er involvert for å hente ut og levere data. Ved tilgang til data på et webgrensesnitt kan altså brukerne av denne dataen hente den ut hvor enn de befinner seg – enten det er på kontoret eller ute i felten. Informant 4 forklarer hvordan dette også gir nye muligheter for å ta vurderinger i felten:

«Det å ha data ute i felten, gjør at du kan gjøre vurderinger i felten også. Vi ønsker at alt skal kunne gjøres via nettbrett som alle ansatte ute har (...).»

Det å ta vurderinger ute i felten er også noe informant 7 legger vekt på:

«I forhold til tidligere så reiste man ut på stasjonen og fulgte en sjekkliste. Nå kan man kontrollere det opp mot reelle data på det man jobber på, for å se at man oppnår de resultatene man vil ha. Det tenker jeg kanskje er en stor forskjell når man har fått dette i Gemini Portal.»

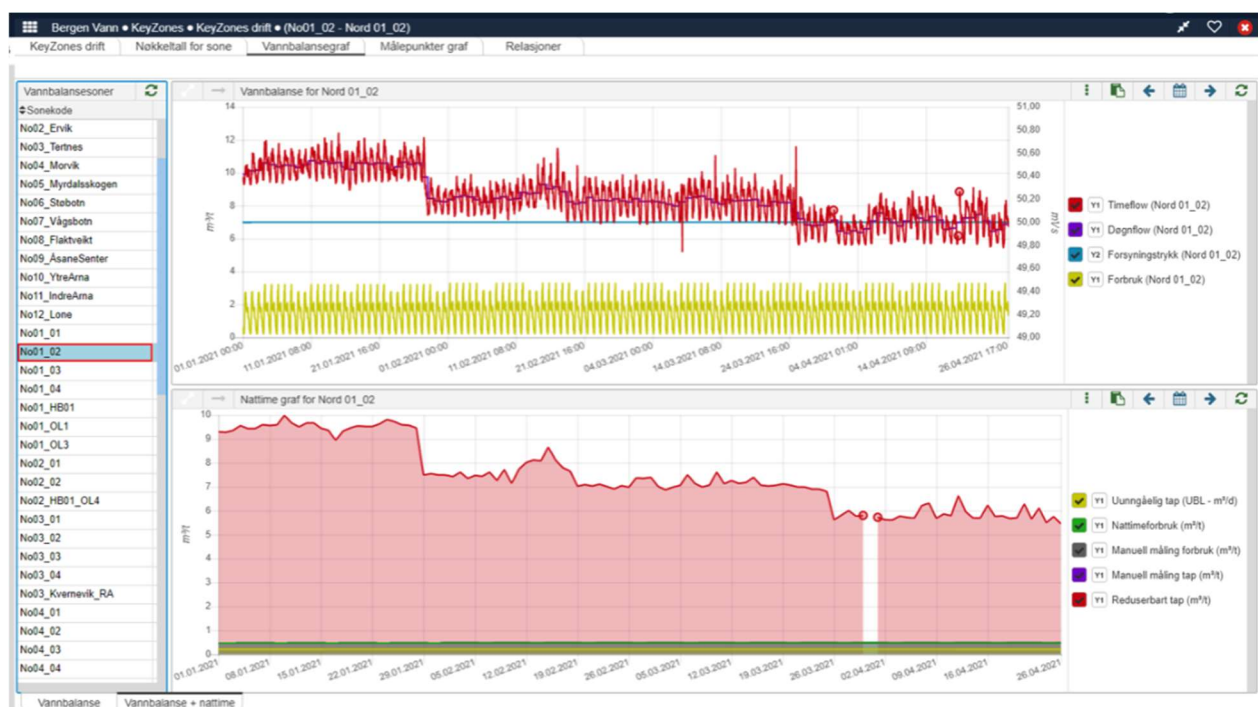
Sitatene belyser at informantene er veldig fornøyd med kartløsningen i Gemini Portal, som omtales som en milepæl i databruk hos VA-etaten. Ved å kombinere oversikt over komponenter i et kart sammen med måledata i sanntid, har VA-etaten fått klarere oversikt og bedre kontroll på nettet. Kartløsningen bidrar også til at måledata blir tilgjengeliggjort for brukerne i felten, noe som medfører at VA-etaten sparer ressurser samtidig som bedre vurderinger kan gjøres direkte i felten.

Funn: Big Data anvendes til å identifisere avvik i vannforbruk

VA-etaten har de siste par årene hatt flere prøveprosjekter med nye løsninger som anvender Big Data til forskjellige formål. En av disse løsningene er lekkasjesøk-verktøyet Keyzones, som nylig er tatt i bruk og er under utprøving av lekkasjesøkere. Tidligere benyttet VA-etaten i liten grad data for å finne lekkasjer, noe som ifølge informant 5 medførte at det gjerne tok flere uker å identifisere en lekkasje. Med tilgangen til Big data blir denne jobben gjort mer

effektivt, hevder informanten. I forlengelse av datainnsamlingen, gav informant 5 en demonstrasjon av hvordan verktøyet Keyzones fungerer.

I Keyzones har man i likhet med Gemini Portal en visualisering av ledningsnett på et kart. Videre er nettet delt inn i mindre geografiske områder, såkalte DMA (District Meter Area). Et DMA er et område hvor alle komponentene i området har vann fra samme kilde. I disse områdene er det plassert ut vannmålere som i sanntid måler hvor mye vann som kommer inn i området og hvor mye som går ut. Basert på data om antall beboere i området og standardtall for normalt forbruk per beboer, kan man gjøre statistiske beregninger i systemet på hva som er estimert normalt forbruk i det aktuelle området. Ved å identifisere avvik mellom hvor mye vann som kommer inn kontra hvor mye som går ut, korrigert for estimert normalt vannforbruk, kan man få bedre oversikt over hvor det er lekkasje på nettet. Dersom det kommer mer vann inn i et område enn hva som går ut, er det et tydelig signal på at det er lekkasje i området. Ved å visualisere målte verdier i sanntid, kombinert med forventede verdier for normalt forbruk, kan lekkasjearbeidere på en tilsynelatende enkel måte identifisere avvik. I figur 5.2 ser vi en skjermdump fra verktøyet Keyzones.



Figur 5.2: Visualisering av avvik mellom målt vannforbruk og forventet forbruk i lekkasjesøk-verktøyet Keyzones (VA-etaten)

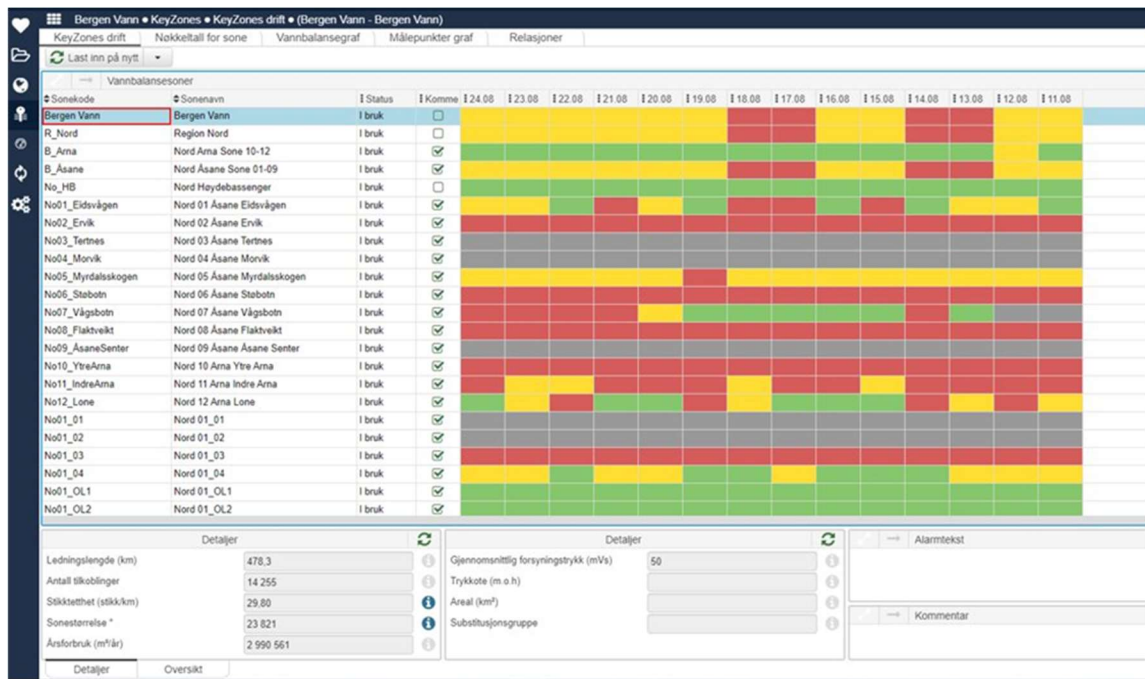
I denne figuren ser vi to grafer, hvor vi skal konsentrere oss om den nederste grafen. Den vertikale aksene viser kubikkmeter vannforbruk per time, mens den horisontale aksene viser tiden i perioden 01.01.2021 til 26.04.2021. Den gule linjen viser et såkalt uunngåelig tap, kort fortalt en liten dråpelekkasje som ikke er økonomisk hensiktsmessig å reparere. Den grønne linjen viser estimert normalt forbruk, altså hva forbruket burde ligge på uten lekkasje eller tap av vann. Den røde linjen viser det faktisk målte vannforbruket i området. Grafen viser her tydelig at det faktisk målte forbruket er mye større enn det estimerte normale forbruket. Vi kan dermed raskt forstå at det er lekkasje ett eller flere steder innenfor dette geografiske området. Denne visualiseringen kan gi lekkasjesøkerne en rask oversikt over vannforbruket i et område, og derav bidra til å snevre inn hvor lekkasjen kan finne sted.

Videre påpeker informant 5 at det jobbes med å montere enda flere vannmengdemålere langs nettet. Dette bidrar til å gjøre områdene enda mindre, som igjen bidrar til at det blir enklere å lokalisere lekkasjene. Informant 5 forklarer hvordan dette hjelper dem å ha god kontroll:

“Jo mindre område man har, jo bedre overvåkning får man (...) Dette systemet viser oss hvor mye vann som er brukt i et område, hvor mye vann vi kan spare, og om vi kommer til å fortsette med det arbeidet. Tidligere så var det normalt å gå ut, reparere lekkasjen og gi oss. Da hadde vi gjort jobben. Men da kunne vi ikke se hvor mye vi kunne spart hvis vi hadde litt mer. Så dette systemet hjelper oss egentlig å ha superkontroll.”

Informanten hevder videre at en av de store fordelene med dette systemet er at det gir tilgang til historiske måleverdier fra sensorene, slik at man enklere kan sammenligne forbruket over tid. Da kan man lettere gjøre korrigeringer for sesongmessige variasjoner i vannforbruket, og derav få en enda bedre kontroll over nettet.

Verktøyet tilbyr også andre former for visualisering, som for eksempel en oversikt over de ulike områdene og en tilhørende fargekode som sier noe om hvor store avvik det er mellom estimerte verdier og faktiske måleverdier. Dette bidrar med å gi lekkasjesøkerne et raskt oversiktsbilde over status på vannforbruk i de ulike områdene, noe som kan være hjelpsomt i prioriteringen av hvilke områder de skal fokusere på.



Figur 5.3: Visualisering av geografiske områder med tilhørende fargekoder for størrelse på avvik (VA-etaten)

Figur 5.3 viser fargekodefunksjonen i Keyzones, med de ulike områdene til venstre, og flere kolonner til høyre med fargekoder for datoer i perioden 11.08.2020 til 24.08.2020. Grønn kode betyr at avviket mellom estimert verdi og faktisk måleverdi er innenfor en godkjent grense. Gul kode betyr at avviket er mer i grenseland, mens rød kode betyr at avviket er såpass stort at det burde prioriteres høyt. Denne funksjonen i Keyzones gir lekkasjesøkere en mulighet til å raskt identifisere hvilke områder som er mest kritiske og som derav burde prioriteres. Informanten bekrefter at ved slike visualiseringer av avvik kan man spare både tid og ressurser i arbeidet med lekkasjesøk.

Funn: Big Data anvendes i prøveprosjekt med maskinlæring til å overvåke og forstå variasjoner i data

Fremveksten av Big Data har gitt nye muligheter for å bruke avanserte analyseverktøy slik som maskinlæring. Big Data kjennetegnes som nevnt av stor variasjon i type data, og mye av dette oppstår fra eksterne kilder utenfor organisasjonen, som for eksempel værdata eller trafikkdata. En av de sentrale motivasjonene bak VA-etatens samarbeid med datasjøen, var å få tilgang til slike eksterne datakilder og i tillegg ta vare på detaljerte måledata fra sensorene slik at det kunne kjøres maskinlæring og andre avanserte analyser på disse dataene. Informant

2 forteller om to driftsområder hvor maskinlæring er tatt i bruk. Det første prosjektet handler om å finne variasjon i vannkvalitetsanalyser. Informant 2 forklarer dette på følgende måte:

«Maskinlæring er på vei full fart inn. (...). Vi tar blant annet enorme mengder med vannkvalitetsanalyser som sendes til laboratorium, som igjen gir data. Vi ønsker å forstå variasjon i de dataene. Men variasjon i vannkvalitet, spesielt med tanke på at vi blander vann fra forskjellige vannkilder og at dette varierer i løpet av året, gir oss en ganske kompleks sammensetning med data. Der jobber vi med maskinlæring for å predikere hvordan vannkvaliteten vil kunne utvikle seg fremover, gitt sånne og sånne variasjoner.»

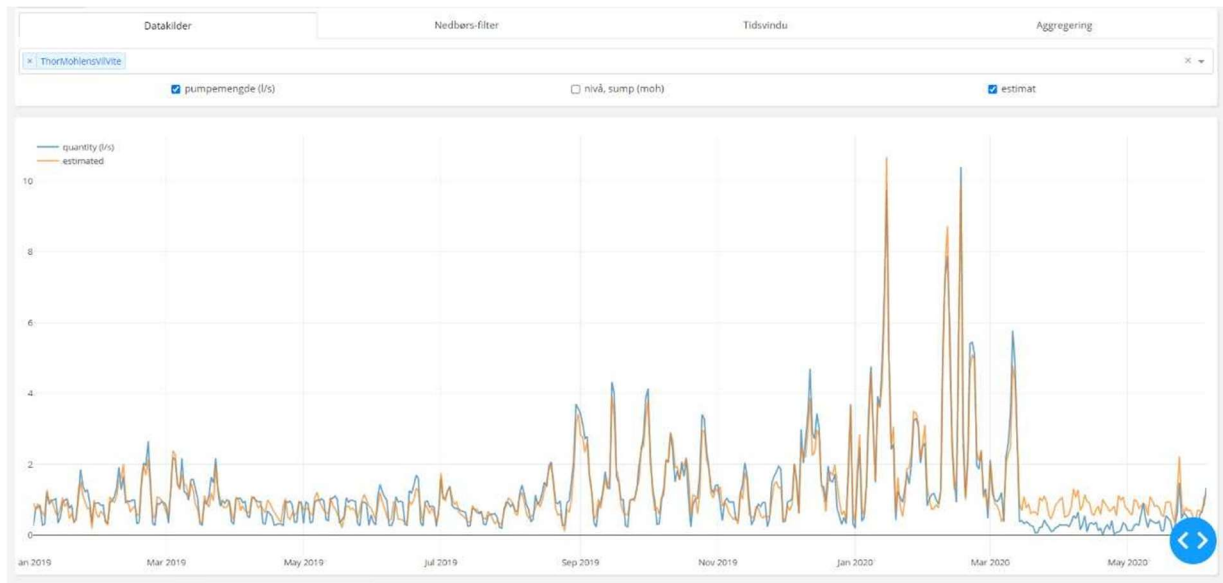
Informanten belyser her at sammensetningen av varierte datakilder gir et komplekst datasett å jobbe med. Maskinlæring vil derfor kunne være et nyttig verktøy for å forstå hvordan ulike faktorer påvirker vannkvaliteten. Informanten forteller videre om sine forventninger til maskinlæring på denne prosessen:

«Maskinlæring vil kunne gi oss en innsikt som vi ikke tidligere har klart, for det har vært for mye data og for mange faktorer til at enkeltpersoner har klart å se den oversikten. Altså, vi har kun én person i Bergen i dag som jobber med vannkvalitet. Og det er klart det er begrenset hvor mye vedkommende klarer å ha oversikt over. Så maskinlæring har jeg absolutt tro på vil være et viktig bidrag.»

Overvåking og kontroll kan få en dypere dimensjon, fordi maskinlæring bidrar til å forstå *variasjon* i driftsdata og hva som påvirker ytelsen til forskjellige driftsprosesser. I stedet for å kun overvåke de faktiske verdiene til vannkvaliteten, kan prediksjoner av hvordan vannkvaliteten burde være gitt de gjeldende forutsetningene, føre til at man kan sammenligne den predikerte verdien mot den faktiske verdien. Dette vil kunne hjelpe med å identifisere avvik og forklare om de faktiske verdiene er slik modellen predikerer at de burde være. Dersom det er stort avvik mellom de predikerte verdiene og de faktisk målte verdiene, kan det være en indikasjon på at noe er feil – enten ved de faktiske verdiene, eller ved modellen.

Et annet område hvor VA-etaten jobber med å få implementert maskinlæring er på avløpspumpestasjoner innenfor seksjonen avløpstransport. Her ble det gjennomført et prøveprosjekt våren 2020 hvor de ønsket å predikere måleverdier på avløpspumper gitt variasjon i ulike variabler, som for eksempel vannmengde, nedbør, tidevann og tid på døgnet.

For å få til dette trengte de å sammenstille store mengder data fra både interne kilder om vannmengder, samt eksterne kilder som nedbørsdata og data om havnivå. Ved å kombinere dette i en maskinlæringsmodell, kunne man få ut en predikert verdi som indikerte hvor mye vann avløpspumpen burde pumpe gitt ulike variasjoner i variablene. Ved å hente ut de predikerte verdiene og legge dem sammen med faktiske måleverdier, fikk man fremstilt en kurve hvor man kunne sammenligne utfallene.



Figur 5.4: Kurver fra en maskinlæringsmodell på en avløpspumpestasjon (VA-etaten)

Figur 5.4 viser kurver over hvor stor mengde avløpsvann som blir pumpet fra en avløpspumpe fra perioden januar 2019 til mai 2020. De blå kurvene viser den faktisk målte verdien fra pumpen, mens de oransje kurvene viser den predikerte verdien som var forventet basert på maskinlæringsmodellen. På den vertikale akse ser vi antall liter som pumpes per sekund, mens den horisontale akse viser tidspunktet. Grafen viser her at det er nokså stor grad av samsvar mellom den predikerte verdien og den faktiske verdien. Dette indikerer at modellen fungerer bra, og kan videre brukes for å overvåke avvik. Informant 6 virket å være godt fornøyd med resultatet fra modellen:

«Nå fungerte det jo veldig bra på denne stasjonen, og det viser jo at dette har noe for seg. Så når man får store endringer i tilrenningen, så kan man faktisk bruke en datamodell som bruker Big data fra både nedbør, tidevann og historiske pumpedata, til å lage en modell som kan sies å stemme med virkeligheten. For det er litt artig at den modellen kan predikere det

her kun basert på nedbør og tidevannet og alt sånt, også stemmer det såpass bra med virkeligheten.»

Samtidig ser vi at perioden mars til mai 2020 har store avvik. Informant 6 forteller at dette kunne indikert lekkasje eller andre kritiske feil med nettet, men informerer om at grunnen til det store avviket skyldes en endring i abonnementenes vannforbruk. Den aktuelle pumpen som ble brukt i prøveprosjektet var lokalisert i nærheten av aktivitetssentre og kontorbygg, som grunnet koronapandemien stengte ned og derav hadde drastisk mindre vannforbruk. Dette er en type variabel som er vanskelig å forutse i en slik modell og kan antyde at man også bør bruke fornuft og intuisjon ved analyse basert på slike modelldata.

På tross av flere suksessfulle prøveprosjekter, påpeker informant 6 at dette ikke er noe som de per nå har fått implementert som en del av den daglige driften:

«I dag så har vi ingen måte å fange dette opp på. Hvis vi hadde slått av den predikerte verdien her, så hadde det ikke vært godt å vite om alt er som det skal være eller ikke. Har det skjedd noe feil der? Jeg vet ikke. Jeg ser at det har skjedd en endring i målingene, men det kan jo skyldes flere ting. For eksempel at det har vært lite regn eller lavt havnivå.»

Informanten tror derimot at dette er noe som kommer til å prege driften mer fremover og at man ved tilgjengeliggjøring av mer data kan få bedre overvåking av nettet:

«Bouvet [konsulentbedrift] ser på å få det opp i datasjøen og få det i produksjon. Da blir modellen oppdatert løpende og vi kan bruke det til å overvåke daglig at nettet fungerer som det skal, eller om det er for lite eller for mye i forhold til hva det burde være. Så dette er helt klart målet å få til.»

Basert på sitatene, kan Big Data gi store muligheter for å drive maskinlæring og avansert analyse i VA-etaten. Maskinlæring bruker Big Data til å predikere variasjoner i driftsyttelse basert på forskjellige variabler, noe informantene hevder gir resultater som 'stemmer med virkeligheten'. På tross av at VA-etaten ikke helt har fått implementert disse løsningene i driften enda, virker informantene å ha store forventninger til at disse løsningene vil gi bedre overvåking og kontroll i fremtiden.

Oppsummering av overvåking og kontroll

Basert på funnene virker Big Data å anvendes til overvåking og kontroll på flere måter. Sensordata i sanntid fra SCADA-systemet anses å være 'synet' til VA-etaten, og avhengigheten til teknologi er tilsynelatende stor ettersom store deler av infrastrukturen ligger under bakken. Implementeringen av kartverktøyet Gemini Portal blir beskrevet som en 'milepæl' i VA-etaten, hvor kartdata er kombinert med måledata i sanntid for å kunne ha oversikt og kontroll over komponenter i felten. Videre brukes Big Data i arbeidet med lekkasjesøk for å identifisere avvik som indikerer lekkasje, noe som anses å gi 'superkontroll' over prosessen. Samtidig brukes Big Data til prøveprosjekt med maskinlæring for å predikere variasjoner i driftsyttelse basert på ulike variabler, noe informantene hevder har gitt resultater som i stor grad 'stemmer med virkeligheten'. På tross av flere suksessfulle prøveprosjekter, er analyse med maskinlæring per nå fortsatt ikke implementert og tatt i bruk, men forventningene er store til at dette kommer til å prege driften stort i årene fremover.

5.2 Anvendelse av Big Data til planlegging

Planleggingsfunksjonen er tett knyttet opp til andre styringsfunksjoner, spesielt overvåking og kontroll, ettersom informasjonen som samles inn her gjerne skaffer grunnlaget til å planlegge hvilke tiltak man skal gjøre videre. Basert på intervjuene innebærer planlegging i VA-etaten mye forskjellig og varierer ut fra hvilken seksjon det er snakk om. Vi har gjennom intervjuene fått innblikk i to planleggingsprosesser; Planlegging av tilsyn og vedlikeholdsaktiviteter, og planlegging rundt utskiftning av vann- og avløpsledninger.

Funn: Big Data anvendes til å utføre behovsbasert planlegging

Basert på funn fra overvåking og kontroll så vi at bruken av sensordata i sanntid kan være en avgjørende teknologi for å holde god oversikt over ledningsnett. I stedet for å overvåke basert på menneskelige antagelser om hvordan tilstanden på ledningsnett er, kan man ved hjelp av sensordata i sanntid overvåke måleverdier for å ha bedre kontroll på hva som skjer i nettet til enhver tid. Flere av informantene hevder at tilgang på denne dataen fører til at planleggingen knyttet til enkelte driftstiltak endrer seg. Hvor etaten tidligere har gjennomført regelmessige rutinekontroller på de ulike komponentene i nettet, vil sanntidsdata fra sensorene kunne føre til at man kan rykke ut der det er behov. Informant 1a forklarer dette:

«På avløp [seksjon] følger operatørene SCADA-systemet. Der får de meldinger om avvik, altså om noe ligger utenfor en normalgrense, også rykker de ut på det. Tidligere så kjørte de runden rundt for å sjekke at alt var som det skulle. Da hadde de faste runder som de kjørte, uavhengig av om det var noe hendelser eller ikke, for da hadde vi ikke noen annen måte å gjøre det på. (...) Så vi har gått fra å kjøre rundt å se på alle komponentene, til å nå få inn data og heller rykke ut på det som er viktig, altså der det er behov.»

Informanten hevder altså at tilgang til sanntidsdata gjør at planleggingen endres fra rutinebasert til mer behovsbasert. I stedet for å gjøre manuelle tilsyn på alt, muliggjør data at man kan rykke ut der det er størst behov. Informant 7 presiserer videre at det i Bergen er totalt rundt 200 pumpestasjoner, hvor det på avløpspumpestasjonene er 2-3 pumper på hver stasjon. Å rykke ut til hver eneste stasjon for å gjennomføre tilsyn krever mye ressurser, noe informantene hevder er unødvendig dersom noen pumper viser seg å være i god tilstand. Da kan man heller sette inn ressursene på de pumpene som faktisk trenger vedlikehold.

Maskinlæring er et verktøy som kan ha stor innvirkning på overvåking og kontroll, men ifølge flere av informantene er dette også noe som kan brukes til behovsbasert planlegging. Informant 6 forteller hvordan innsikt i variabler som påvirker vannmengden i pumpestasjonene kan brukes til å sette i gang nødvendige tiltak *før* en hendelse inntreffer:

«Nedbørsdata, havnivå og de historiske sensordataene kan brukes til å lage en prediksjon frem i tid om hva og hvilke mengder som kommer til å renne inn i stasjonen. Dette gir innsikt som kan brukes til å sette i gang tiltak hvis det er en stasjon som ikke har nok kapasitet til å pumpe inn alt. Så ved for eksempel ekstremnedbør og stormflo så kan man bruke denne innsikten til å sette inn tiltak før hendelsen inntreffer.»

Maskinlæring basert på Big Data kan ha potensiale til å predikere hendelser frem i tid, slik at nødvendige tiltak kan planlegges og iverksettes før ulike hendelser skjer. Dette er ikke bare en fordel for overvåking og kontroll, men kan også gi mer ressurseffektiv planlegging ved at man setter inn de nødvendige tiltakene der behovet er størst.

Samtidig som flere informanter har stor tro på at Big Data vil ha stor innvirkning på planleggingsfunksjonen, er det enkelte som mener at det per i dag fortsatt gjennomføres mye rutinekontroller og tilsyn. Informant 6 forteller dette:

«Nå er det mer sånn tut og kjør når det skjer noe akutt, også går vi på mange tilsyn som strengt tatt ikke er nødvendig å gå på. Nå tar vi alle stasjonene og pumpene opp på en årlig sjekk. Målet er jo at vi skal spare oss for alle de tilsynene. Istedenfor å bare gå ut på rutinebesøk og gjennomføre rutinekontroller, så tar vi det etter behov. Med Big Data kan vi se hvilke stasjoner som sliter, eller har en pumpe som er dårlig, og derfra går vi inn å gjør tiltak. Så det er det vi ønsker da, at det blir mer behovsbasert drift.»

Basert på denne informantens uttalelser kan det virke som at det er litt ulike oppfattelser av hvor mye tilsyn og rutinebesøk som gjennomføres i dag. Det er derfor litt usikkert akkurat hvor langt VA-etaten har kommet med databruk på dette området, og det er ikke utenkelig at det kan være visse variasjoner mellom de ulike seksjonene. Informant 7 påpeker at dataen er tilgjengelig, men at den kanskje ikke utnyttes fullt ut enda:

«Altså, pumpe-verden i dag har kommet så langt at pumpene kan fortelle mye mer selv til et system om hvilken tilstand de har; om de jobber lett eller tungt, eller om de leverer det de skal eller ikke. Så jeg tenker vi må bruke den dataen der til bedre driftsplanlegging da.»

Hvorfor det ikke utnyttes bedre kan være mange grunner til, men informant 7 hevder at det kanskje kan være litt datavegring og frykt for store endringer i organisasjonen:

«Det er jo kanskje litt sånn i en offentlig organisasjon, at det er litt frykt for store endringer, altså det sitter litt i veggene det der. Folk kan jo si; «Ja.. nei, det var jo enklere før når vi bare kjørte ut og så komponentene fysisk.». Men hvis du tenker litt økonomi og drift oppi det her, så er det jo en fordel hvis du kan ha en bedre vedlikeholdsplan som du faktisk kan forholde deg til, og som bedre utnytter de ressursene du har.»

Datavegring generelt er en potensiell utfordring flere av informantene legger vekt på i intervjuene. Dersom ansatte ikke tar i bruk nye digitale løsninger, får man heller ikke realisert den potensielle verdien knyttet til disse løsningene. Basert på dette sitatet er det derfor avgjørende å dyrke god digital arbeidskultur med god opp trening i bruk av digitale verktøy.

Sitatene fremmer en tydelig forventning om at Big Data vil føre til et skifte fra rutinebasert planlegging til mer behovsbasert planlegging. Særlig tilgangen til sanntidsdata, samt sammenstilling av informasjon fra flere kilder, er en sentral driver for dette skiftet. Datavegring og frykt for store endringer kan derimot være en potensiell trussel for å få realisert verdier fra behovsbasert planlegging.

Funn: Big data brukes lite til planlegging rundt utskiftning av rør

VA-etaten har et mål om skifte ut rundt 0,7 % av vannledningsnettet, og rundt 1 % av avløpsnettet hvert år. Dette tilsvarer rundt 7-8 km med vannledning og rundt 10 km med avløpsledning i året, ifølge informant 8. Måltallet er basert på forventet levetid på rørene før de er moden for en utskiftning. Informant 8 påpeker at planlegging og prioritering av hvilke ledninger som skal skiftes ut, er en stor og komplisert jobb hvor det må tas hensyn til veldig mange faktorer. Det er stor variasjon i alder og tilstand på ledningene, hvor de tidligste ledningene ble lagt helt tilbake på 1860-tallet. Selv om det er logisk å tenke at man prioriterer utskiftning av de eldste ledningene, er det ikke fullt så enkelt skal vi tro informant 8:

«En ledning fra 1920 kan være mye bedre enn en ledning fra 1960. Grunnet til det kan være at de gjorde en veldig god jobb når de la ledningen fra 1920-tallet, mens på 50-60-tallet etter krigen så begynte de med gravemaskin og større utstyr som gir mer forstyrrelser og fører til at det blir gjort dårligere arbeid.»

Videre er det ikke nødvendigvis slik at VA-etaten kan prioritere og skifte ut akkurat de ledningene som de selv mener burde prioriteres. Ledninger ligger som oftest under bakken, og må derfor sees i sammenheng med annen infrastruktur slik som trafikk og vei. Dette er et komplekst samspill hvor det må tas hensyn til mange samfunnsaktører. Det er gjerne dialog med alt fra Statens Vegvesen til strømbyggere for å sikre seg at man kan gjøre mest mulig arbeid samtidig, uten å måtte gjøre grave opp mer enn nødvendig. På tross av stadig mer data om sitt eget ledningsnett, påstår informant 8 at de eksterne hensynene preger prioriteringene stort:

«I dag er det så mye mer data tilgjengelig enn noen gang tidligere. Dette påvirker også prioriteringene. Men samtidig så er mye av det som styrer våre prioriteringer i dag eksternt, altså utbygginger fra andre etater, slik som vei, bybane, bossnett osv. Så hvis vi bare skulle prioritere ut fra det som vi selv tenker burde vært gjort, så ville det kanskje vært litt andre prioriteringer. For i virkeligheten er det mange andre faktorer som gjør at vi må prioritere annerledes.»

Informant 1a legger videre vekt på at de i utskiftningsprosessen planlegger for et veldig langt perspektiv sammenlignet med andre aktører, noe som gjør at de er nødt til å ha et helhetssyn i planleggingen:

«Det er så mange andre hensyn vi skal ta, og da kommer dette helhetssynet inn. Det handler om å se våre resultat i sammenheng med andre sine behov og ønsker. For det vil jo aldri bli feil om du skifter en vannledning eller en kloakkledning som er dårlig. Men det er ikke sikkert det er akkurat den du hadde mest behov for å ta i dag. Kanskje du kunne ventet i 20-40 år. Vi jobber med anlegg som skal leve i 100 år, mens mange ting rundt oss skal ikke leve så lenge.»

Informant 8 forteller et eksempel hvor de i 2019 hadde planlagt å skifte ut en ledning fra 1926 som lå under bakken langs hovedveien i Sandviken. Grunnet den planlagte utbyggingen av bybanen mot Åsane måtte VA-etaten avvente utskiftningen av røret i tilfelle bybanen hadde planer om å bygge i området innen kort tid. Det hadde vært lite gunstig å starte graving og utskiftning av ledningen dersom bybanen kom et par år senere og gravde opp alt på nytt. VA-etaten måtte derfor endre sine egne prioriteringer for å ta hensyn til andre. I etterkant oppstod det derimot et stort brudd i røret, noe som førte til akutt utrykning og behov for utskiftning. Dette akutte inngrepet førte til at de måtte stenge hele hovedveien i over et døgn, noe som fikk enorme konsekvenser for trafikken og medførte store kostnader for samfunnet. I lys av hendelsen fikk VA-etaten mye kritikk og negativ medieomtale, og informanten hevder at det er vanskelig å få forståelse i media om hvorfor ledningen, som altså var fra 1926, ikke var byttet tidligere. Et slikt vannledningsbrudd er vanskelig å forutse, men får altså enorme konsekvenser.

De siste årene har fokuset innenfor ledningsfornying endret seg til å bli mer konsekvensorientert. Informant 8 forteller om denne prosessen:

«Tidligere ble ledninger med mye brudd prioritert og fornyet først, for når det er brudd på ledninger så tyder det jo på at de er dårlig. Men det som vi har begynt å tenke mer på nå i forhold til før, er konsekvenser. Hvis du har en liten ledning med mye brudd på så er jo det alvorlig nok det, men det er ikke så store konsekvenser. Konsekvensene når en stor ledning får brudd er ofte mye større og vi får mye større følgeskader (...). Så trenden nå er at vi fokuserer på konsekvensene ved brudd og hvor ledningene ligger. Ved bruddet i Sandviken i fjor lå vannledningen midt i hovedveien og gav kjempestore konsekvenser for trafikken.»

Etter hvert som VA-etaten har fått på plass flere vannmålere og bedre sensorer, har datamengdene økt og det er blitt lettere å oppdage ledninger med brudd. I dag registreres ledninger med mye brudd i Gemini-systemet slik at de som driver med utskiftning har oversikt over disse. Men som informanten forteller, må det tas hensyn til både eksterne etater og private for å kunne tenke konsekvenser. Dette gjør arbeidet med å planlegge utskiftninger vesentlig mer komplekst, og fører til at man ikke kun kan basere planleggingen på interne data. Ifølge informanten trengs det mer eksterne data og nye programmer som kan sammenstille flere hensyn. Informanten håper derimot at mer data og bedre dataprogrammer vil gjøre prosessen bedre i fremtiden:

«Vi er på vei til å bli mye bedre på planleggingen av utskiftningen ved at vi får mer data og bedre dataprogram som kan hente inn ulik informasjon som for eksempel trafikkdata og flomveier.»

Ved å kombinere data fra ulike kilder håper informanten at man kan få et bedre bilde av konsekvenser. For eksempel kan trafikkdata i sammenstilling med interne data om bruddledninger bidra til å forstå hvilke konsekvenser som vil inntreffe dersom man ikke skifter en ledningen. Ved å anvende slike data i avanserte analyseverktøy vil det kunne hjelpe å prioritere hvilke ledninger som skal byttes og når. På tross av at dette ikke er utviklet og implementert helt enda, er det flere informanter som nevner at det jobbes aktivt på denne fronten med å teste ut nye løsninger. Et ledningsfornyingsprogram fra leverandøren Powel er en løsning flere håper skal kunne bidra til bedre planlegging rundt utskiftning av nettet fremover.

Sitatene belyser at planlegging og prioritering knyttet til utskiftning av ledninger er en kompleks jobb. Hovedgrunnen til dette er at VA-etaten må ta store hensyn til andre samfunnsaktører i dette arbeidet. Grunnet hensynet til andre aktører er planleggingen blitt mer konsekvensorientert. Enkelte informanter håper at mer data fra eksterne kilder, samt nye analyseverktøy, kan bidra til å gjøre planleggingen bedre i fremtiden.

Oppsummering av planleggingsfunksjonen:

Basert på funnene kan Big Data ha nokså stor innvirkning på planleggingsfunksjonen. Tilgang på sensordata i sanntid bidrar til et skift i planleggingen av tilsyn og vedlikehold, som går fra å være rutinebasert til å bli mer behovsbasert. Samtidig er planlegging knyttet til utskiftning av ledninger i stor grad preget av hensynet til eksterne aktører, noe som har ført til et mer konsekvensorientert fokus i denne prosessen. Hensynet til eksterne aktører ser derimot ut til å gjøre prosessen betydelig mer komplisert. Informantene har store forhåpninger til at data fra eksterne kilder, samt nye analyseverktøy, vil bidra til bedre konsekvensorientert planlegging i fremtiden.

5.3 Anvendelse av Big Data til beslutningstaking

I kapittel 2.2.1 er beslutningstaking tett knyttet opp til de andre styringsfunksjonene ettersom informasjonen som fremkommer i disse funksjonene ofte munner ut i en eller annen form for beslutning. I forhold til planlegging handler prioriteringer om å ta beslutninger om hva man skal velge. Funn fra de andre funksjonene som handler om å ta beslutninger, spesielt planlegging, vil derfor være relevant også for denne funksjonen.

Funn: Big data brukes som informasjonsgrunnlag i beslutningstaking, hvor intuisjon og erfaring fortsatt er viktig

Fra kapittel 5.1 så vi at Big Data har gitt nye muligheter til å få informasjon - både gjennom interne sensordata i sanntid og sammenstilling med eksterne kilder. Dette har bidratt til å få en oversikt og kontroll over ledningsnettene som VA-etaten ikke tidligere har hatt. Anvendelsen av disse store datamengdene inn i diverse verktøy som Keyzones og maskinlæringsmodeller har også bidratt til å kunne se trender og sammenhenger som man tidligere ikke har hatt kjennskap til, noe som igjen har påvirket beslutninger. På spørsmål om hvilken rolle data spiller i beslutningstaking svarer informant 2 følgende:

«Nå er jeg ingeniør da, så jeg liker jo selvfølgelig å ta beslutninger basert på 'fakta' og en analytisk tilnærming. Og i den forbindelse så er jo data den beste fakta du kan ha.»

Informanten avskriver at han tar beslutninger kun basert på data, men denne uttalelsen bekrefter det mange ser på som 'evidensbasert beslutningstaking' (Holsapple et al., 2014) hvor data gjerne er den viktigste kilden i beslutningstaking. Samtlige informanter hevder samtidig at data fungerer som et informasjonsgrunnlag som må sees i sammenheng med erfaring og intuisjon. Informant 7 påstår at en kombinasjon av data og kunnskap er bra for beslutningstaking:

«Jeg tror du må ha en viss forståelse for oppbygging av data og det man jobber med, altså noen knagger å henge dataen på. Jeg tror det blir utfordrende å ta beslutninger basert på ren data, men en kombinasjon er jo fantastisk.»

Informant 3 støtter dette og hevder at data fungerer som et godt informasjonsgrunnlag, men at erfaring er en viktig faktor for å kvalitetssikre informasjonen i dataene:

«Intuisjon og erfaring er viktig for å kontrollere om datagrunnlaget ditt er feil eller ikke. Det er ikke sjelden at man beregner noe, også tenker man at det ikke stemmer med hva man trodde. Så da går man tilbake og finner ut at det var noe feil, for eksempel at folk brukte tommer istedenfor millimeter. Så på det nivået er erfaring en viktig kontroll for å vurdere om tallene som beregnes virker uvirkelige eller ikke. Men samtidig kan man ikke gjøre det på for mange tall, så man er nødt til å støtte seg til data for å optimalisere (...). Data er uansett viktig for å få et godt informasjonsgrunnlag, og for å ta de riktige beslutningene».

Informant 3 forteller videre om et eksempel hvor personer i VA-etaten i for stor grad stolte ukritisk på data i forbindelse med utbygging av et nytt renseanlegg. Ved beregning av dimensjonene benyttet man data som fakta og kontrollerte ikke disse beregningene opp mot intuisjon. Resultatet ble ifølge informanten dårlig bruk av ressurser:

«Når de gjorde beregningene og så på tallene, kunne de faktisk se at den mengden forurensning som ble målt inn var mye større enn det tallene skulle tilsi. Men de valgte å tro at det kanskje var en bedrift eller liknende som var koblet på, noe det ikke var. De brukte data direkte, uten å bruke erfaring eller intuisjon til å kontrollere beregningene (...). De støttet seg kun på data og var ikke kritisk nok til denne. Anlegget har derfor to svære rensebasseng, men i dag bruker vi bare ett av dem. Så nå er det overdimensjonert i lang tid fremover.»

Videre forteller informant 1a at de på ingen måte har et dataverktøy som de kan basere alle sine beslutninger på, og hevder at for å kunne ta hensyn til andre i omgivelsene og bevare et helhetssyn er det viktig at mennesker bruker sine evner:

«Vi har ingen verktøy som vi tar alle beslutninger basert på. Vi er en del av et samfunn og spesielt i urbane områder er det ulike ting som må tas hensyn til og vektet. Men det å ha et verktøy som raskt gir oversikt over eget system er til hjelp, og det at vi har kjørt kontroll på det. (...) Det å se totalen vil alltid kreve litt menneskelig energi og tid før vi går videre til endelig beslutning.»

Informant 8 påpeker at det er viktig med kunnskap og kompetanse, men at det også kan være vanskelig å vite hvor mye man skal vektlegge magesfølelse og erfaring i ulike situasjoner:

«Hvis man putter dårlig data inn, så får man dårlig data ut. Det er det ikke tvil om. Så jeg både tror og håper at vi må ha kunnskaper likevel. En bør ha en kompetanse. Også er det jo vanskelig dette med å svare på hvorfor man gjør de valgene man gjør. For det er jo mye magesfølelse og erfaring, og det er ikke alltid så lett å vite hvordan man skal vektlegge den erfaringen. Også er det ikke så mye fasiter heller.»

Basert på sitatene er informantene samstemte om at Big Data spiller en viktig rolle for å ta beslutninger. Etersom Big Data gir informasjon om nettet i sanntid, får beslutningstakere tilgang til tidsrelevant informasjon som kan danne grunnlag for å ta gode beslutninger. Tross dette, er informantene tydelig på at Big Data fungerer som et informasjonsgrunnlag som fortsatt må sees i sammenheng med erfaring og intuisjon når man skal ta beslutninger.

Funn: Data av høy kvalitet er viktig for å kunne ta gode beslutninger

I forlengelse av hvordan Big Data anvendes i beslutningstaking, fortalte flere av informantene om hva de mener er viktig når man tar beslutninger basert på data. Informant 2 forteller at robust data som man kan stole på er viktig for en driftsorganisasjon som VA-etaten:

«Nå er jo jeg i en driftsorganisasjon, så det er viktig at det er robust, at det virker og at det ikke stadig vekk faller ut slik at vi må restarte og alt mulig sånn dill. Det er alfa og omega i en litt sånn low-tech organisasjon som vi er. Det andre er selvfølgelig at dataen må være korrekt og av høy kvalitet. Det opplever ikke jeg at er det stort problem hos oss, altså. Om det er måling av trykk, temperatur eller mengde, så er det stort sett presist nok for oss.»

Det at dataen er korrekt er noe som støttes av flere av de andre informantene. Dette henger også sammen med det å bruke intuisjon til å kontrollere at informasjonen stemmer, noe informant 3 vektlegger:

«Først må du vite at den dataen du skal bruke er korrekt. Det er kanskje det viktigste. Det er ikke sjelden at den data som folk bruker er feil. Så jeg tror det er viktig at du har en grundig kontroll på at den dataen som du baserer ting på, er korrekt. Også må du være ærlig og ikke

bruke data dersom du ikke er trygg på den og tørre å sløyfe den hvis det ikke stemmer med hva din intuisjon tenker er rett.»

Informant 5 påpeker videre at nøyaktigheten på data er en viktig faktor, og at feil data kan føre til feil tiltak:

«Nøyaktighet er viktig. At man kan stole på data er en veldig viktig faktor. Det finnes så mange produkter med forskjellig nøyaktighet. (...) Så derfor må man velge riktig utstyr og verktøy slik at man minimerer feilkilder. Fordi feil data gir feil resultat og dermed feil tiltak».

På oppfølgingsspørsmål om informantene opplever at dataen VA-etaten benytter i dag er bra, svarer vedkommende:

«Ja. Over 90 %, vil jeg si. Men det er ikke alltid vi har data heller. Vi har noen ganger datautfall, og det er på grunn signal eller at vi har mistet dekning grunnet strømbrydd eller noe sånt. Men generelt sett så har vi god data, og vi har god oppetid i disse systemene.»

Sitatene belyser at data av høy kvalitet, som er korrekt og nøyaktig, er avgjørende for å ta gode og riktige beslutninger. Generelt er informantene tilfreds med kvaliteten på dataene som benyttes i VA-etaten.

Oppsummering av beslutningstaking

Basert på sitatene er informantene samstemte om at Big Data spiller en viktig rolle for å ta beslutninger. Informasjon om nettet i sanntid, gir beslutningstakere tidsrelevant informasjon som danner et godt grunnlag for å ta gode beslutninger. Likevel er intuisjon og erfaring fortsatt viktige faktorer i beslutningsprosessen, som blant annet bidrar til å kontrollere dataene. I tillegg til dette beskriver sitatene at data må være av høy kvalitet, altså korrekt og nøyaktig, for å kunne ta gode og riktige beslutninger. Generelt er informantene tilfreds med kvaliteten på dataene som benyttes i VA-etaten.

5.4 Anvendelse av Big Data til resultatmåling

I denne oppgaven har vi valgt å konsentrere oss om resultatmåling som prosessen med å måle KPI'er, *Key Performance Indicators*. I datainnsamlingen lot vi informantene selv legge vekt på hvilke KPI'er de ville fortelle om.

Funn: Måledata fra styrings- og overvåkingssystemet brukes til resultatmåling

I intervjuene observerte vi at de ulike seksjonene fokuserer på ulike KPI'er. Avløpsrens har for eksempel KPI'er knyttet til kostnader for rensing av avløpsvann, mens vanddistribusjon blant annet måler totalt vannforbruk. Informantene forteller at forskjellige produksjonstall og mengdetall hentes fra styringssystemet SCADA, mens økonomitall hentes fra etatens økonomisystem, Agresso. Måten man regner seg frem til de ulike KPI'ene varierer, hvor noen kun er basert på måleverdier fra organisasjonens styringssystem, mens andre krever mer bearbeiding og sammenstilling med annen data. SCADA-systemet virker imidlertid å være sentralt i de fleste KPI'er og anses som et viktig verktøy. Informant 3 forklarer:

“Vi kan ikke ha KPI'er uten å ha samlet all dataen som vi har, spesielt gjennom styringssystemet som er det viktigste verktøyet for oss.”

Sitatet påpeker at man ikke kan måle KPI'er uten data. Måleverdier i sanntid fra SCADA-systemet er avgjørende for å kunne måle hvordan ulike seksjoner leverer på sine måltall. Når disse verdiene samtidig lagres historisk, vil det kunne bidra til å måle resultater over tid. Tilgang til detaljerte måledata kan derfor gi mulighet for gode resultatmålinger.

Funn: Prosessen med å måle enkelte KPI'er preges av manuelt arbeid som følge av at systemer ikke snakker sammen

Mens det er relativt enkelt å hente ut måledata for totalt vannforbruk fra SCADA-systemet, krever enkelte KPI'er at man henter data fra ulike systemer og sammenstiller dette manuelt. For eksempel måles kostnaden per kubikkmeter produsert biometan i avløpsrens-prosessen, hvor det blir hentet ut data fra både SCADA og økonomisystemet. Informant 3 påpeker imidlertid at disse systemene ikke er koblet sammen og at dataen dermed må sammenstilles manuelt. Manglende integrasjon blant de ulike systemene virker å være et gjentakende tema, og andre informanter ønsker også at systemene gjerne skulle snakket bedre sammen. Informant 3 forklarer:

“På det øverste nivået her jobber jeg med at vi skal ha fem mål eller KPI'er hvor vi er nødt til å hente data fra flere systemer og akkurat nå må jeg gjøre deler av det manuelt.”

Videre påpekes det at dette er en av grunnene til at noen av KPI'ene ikke måles kontinuerlig og kun oppdateres en sjelden gang. KPI'er som for eksempel klimaavtrykk beregnes kun årlig på grunn av denne problemstillingen, og informanten 3 uttrykker ønske om et mer automatisert system slik at KPI'er kan måles mer regelmessig:

“Akkurat nå beregner jeg bare de tallene én gang per år fordi jeg må kombinere de forskjellige datakildene manuelt, og det er ikke verdt å gjøre det løpende hver måned. Det ville vært kjekt å ha de løpende per måned, men vi har ikke et system som kan det enda.”

Basert på sitatene er enkelte KPI'er utfordrende å måle ettersom de krever sammenstilling av data fra ulike systemer som ikke snakker sammen. Dette fører til at disse KPI'ene gjerne måles på en sjeldnere basis enn ønsket, fordi det av hensyn til tid og ressurser ikke er hensiktsmessig med mer regelmessige målinger.

Funn: Tilgang på data fra flere kilder kan gi mer presise målinger av KPI

Enkelte av informantene fortalte at de gjerne skulle hatt mer tilgang til data fra andre kilder og at dette kunne bidratt til mer presise målinger av KPI. Informant 6 peker på at Big Data kan bidra med nedbørsdata og tidevannsdata som kan benyttes til å se på endringer i hvor mye fremmedvann som kommer inn fra ulike områder på nettet. Tilgang på data fra eksterne kilder kan på denne måten bidra til å forstå hvor mye fremmedvann som kommer inn, og derfor kunne gi mer presise målinger. Informanten presiserer derimot at de ikke er helt i mål med slike typer målinger enda. Informant 4 forteller videre om et ønske om å se værdata inn i systemet for å kunne gi en bedre forståelse av tallene:

«Vi ser måned for måned og sammenligner med året før, eller utviklingen fra de tidligere årene. Så vi prøver jo, men det er mange faktorer som gir endring i vannforbruk. Været kan jo være en variabel som bør inn i det systemet, men i dag er ikke dette en parameter som vises særlig godt. (...) Så det å ha temperatur og klimatall som kommer inn i systemet gir en bedre forståelse av tallene. Nå er det graving manuelt i ettertid for å forstå endringene.»

Tilgang på data fra eksterne kilder, slik som værdata, er noe VA-etaten jobber med å få bedre integrert i systemene. Målet er at dette kan gi bedre forståelse av tallene og mer presise målinger av KPI. Sett i sammenheng med forrige funn virker det derimot å være behov for nye systemer som er bedre integrert for å få til dette.

Oppsummering av resultatmåling

Basert på funnene virker måledata i sanntid fra SCADA-systemet å være helt avgjørende for å kunne måle KPI'er i VA-etaten. Samtidig indikerer funnene at prosessen med å hente ut enkelte KPI'er preges av mye manuelt arbeid ettersom systemer er dårlig integrert. På en annen side håper enkelte informanter at tilgang på eksterne datakilder som vær og klima kan bidra til å gi mer presise målinger av KPI i fremtiden. Dette forutsetter derimot nye systemer som er bedre integrert.

5.5 Anvendelse av Big Data til rapportering

Rapportering i VA-etaten forekommer både som intern rapportering innad i organisasjonen, samt ekstern rapportering til myndigheter. Ekstern rapportering til myndigheter er kanskje spesielt viktig, ettersom VA-etaten er en offentlig driftsorganisasjon med forvaltningsansvar for en samfunnskritisk funksjon.

Funn: Prosessen med rapportering er preget av mye manuelt arbeid

Basert på intervjuene foregår mye av rapporteringen i VA-etaten nokså manuelt. I flere av tilfellene henter driftsledere måledata fra styringssystemet, som de manuelt flytter inn i et Excel-skjema. Dette sendes så til seksjonsleder som videre aggregere opp tallene. Informant 3 forteller om denne prosessen innenfor seksjonen avløpsrens:

«På alle renseanleggene genererer driftsleder hver måned en rapport som inneholder data som vannmengde, vann inn og ut, innsatsmidler slik som mengde kjemikalier brukt, energiforbruk osv. Driftsledere genererer en rapport hver uke som vi trykker inn i et Excel-ark. Så da har jeg 6 av disse Excel-arkene som jeg kan aggregere opp for rapportering.»

Hvorvidt prosessen med rapportering i de ulike seksjonene er standardisert eller ikke, er uvisst. Dersom den ikke er det, vil det kunne gå med mye ressurser til å korrigere og endre på struktur for å se at ting henger sammen. På spørsmål om hvorvidt økonomiske tall rapporteres automatisk eller ikke, svarer informant 4 følgende:

“Det er et uttrekk fra datagrunnlaget, men jeg gjør en del manuell korreksjon eller kontroller for å få det til å henge sammen.”

Basert på denne informantens forklaring, kan det tyde på at manuelle korrigeringer er nødvendig og prosessen kanskje burde vært mer standardisert. På videre spørsmål om hva informant 4 tror automatiserte rapporteringsverktøy kan bidra med, svarer vedkommende:

«Det er jo for å ikke dobbelt registrere eller rapportere, og prøve å få den til å samkjøre. Så når du trenger en rapport så skal den kunne hjelpe til med å være en standard slik at man slipper å gjøre korreksjoner hele tiden.»

Hvorvidt informanten selv mener det er for mye manuelt arbeid med dagens rapportering, kommer ikke tydelig frem, men forventningene vedrørende automatiserte rapporteringsverktøy kan antyde at informanten gjerne skulle sluppet mye av dagens manuelle rapporteringsprosesser. Sitatene belyser uansett at prosessen med å utarbeide rapporter er nokså manuell.

Funn: Tilgang til detaljert rådata kan automatisere generering av rapporter

Fra innledende møter med VA-etaten og ansatte fra datasjøen, kom det frem at VA-etaten har et ønske om å automatisere produksjon av enkelte rapporter som tradisjonelt utarbeides manuelt av de ansatte. I ett av disse innledende møtene, fortalte representanter fra datasjøen om prosessen knyttet til automatisering av rapportene. Dette danner hovedgrunnlag for funnet.

Sitater fra forrige funn forklarte hvordan driftsledere og andre ansatte legger inn store mengder verdier fra diverse måleparametere inn i et stort Excel-skjema. Basert på dette skjemaet lager de en rapport som sendes til sentrale myndigheter. Prosessen med å manuelt legge inn de ulike verdiene virker å være ressurskrevende, både med hensyn til tid og hvor mange folk som er involvert i prosessen. Ved manuell utarbeidelse kan også risikoen øke for at ansatte gjør feil når de skal legge inn tall i Excel-skjemaet. For å kvitte seg med dette store regnearket ønsker VA-etaten å benytte datasjøen sitt historiske Big Data-lager med sensordata i sanntid. Basert på denne datamengden er ønsket til VA-etaten å sette opp en dataflyt inn i et rapporteringsverktøy som automatisk kan generere forskjellige rapporter.

VA-etaten har på dette området nylig jobbet tett med team fra datasjøen for å automatisere ulike rapporter for alle VA-seksjonene. Verktøyet som ble valgt til dette er Microsoft Power BI – et rapporteringsverktøy med et enkelt brukergrensesnitt og muligheter for visualiseringer gjennom grafer og diagrammer. Dataflyten inn i dette verktøyet hentes fra datasjøen. Ifølge en av de ansatte i datasjøen er det flere fordeler med å hente dataflyt herfra. For det første er kapasiteten i datasjøen mye større enn i fagsystemet, noe som gjør at de kan lagre historiske sanntidsverdier langt tilbake i tid. Ved å ha detaljerte historiske data lagret i datasjøen, gis det muligheter for å sende dataflyt inn i rapporteringsverktøy som kan automatisere generering av rapporter. For det andre hevder vedkommende at all data i sjøen er lagret som rådata, altså i sin naturlige form. Dette gjør det enklere å transformere datamengden slik man ønsker, alt etter hvilket behov man skulle ha for dataen.

En av de store utfordringene som utvikleren i datasjøen påpeker når den skal utvikle Power BI rapportene, er å få ned datamengden som går inn i programvaren. Sensormålinger som registreres hvert minutt gir ekstremt mange historiske verdier over tid. For å få ned datamengden gjøres det derfor aggregeringer, slik som å benytte snittverdi av timesverdi for hvert minutt i rapporten. Uten slike aggregeringer, blir datamengden for omfattende til at verktøyet klarer å håndtere det.

Funn: Prosessen med å utvikle gode rapporteringsverktøy krever god kommunikasjon og samhandling mellom utviklere og brukere

Prosessen knyttet til utvikling av nye verktøy, er et sentralt aspekt i forbindelse med å studere hvordan VA-etaten arbeider med Big Data. Måten arbeidet med å automatisere rapportene er organisert på i dag, er ved at personer fra datasjøen og fagpersoner fra VA-etaten møtes i tverrfaglige team. Her skal behovene og ønskene fra fagpersonene legges frem, slik at utviklerne kan få på plass rapporter med nytteverdi for brukerne. I startfasen er det team fra datasjøen som drifter selve Power BI rapportene, men ønsket på sikt er at fagpersonene i VA-etaten skal kunne utvikle og ta over denne delen av jobben også.

På tross av at de fleste informantene hadde vært introdusert eller vært på kursing i Power BI, var det ingen av disse som kunne fortelle at de faktisk hadde implementert verktøyet eller tatt i bruk rapportene herfra. Det kan være flere grunner til hvorfor, blant annet at verktøyet kanskje fortsatt er i en implementeringsfase og at de ansatte derfor ikke har fått tatt det skikkelig i bruk. Samtidig kan det være at det valgte verktøyet ikke er like godt egnet for alle seksjoner og de operasjonelle prosessene her. Informant 3 var en av dem som ikke virket å være fornøyd med Power BI løsningen:

«Power-BI er “business graphics” og ikke et vitenskapelig verktøy. Så jeg tror at vi fortsatt må skaffe et annet program, for eksempel noe som heter MATLAB – som er et avansert program for statistisk behandling av data, som også kan hente data fra datasjøen. For i renseanlegget er det mange variabler og jeg tenker at vi må ha andre typer analyseverktøy enn det som er tilgjengelig via et slikt “business graphics” program.»

Dette belyser at kanskje ikke alle har like godt nytte av det spesifikke verktøyet og av den grunn hadde sett for seg andre løsninger. Informant 2 virker heller ikke å ha tatt verktøyet i bruk:

«Jeg har jo vært på kurs, men så brukte jeg det ikke noe videre, så nå har jeg egentlig glemt det litt ut da. Men han som er økonomisjef hos oss, var på det samme kurset og begynte å bruke det med en gang til å lage økonomirapporter(..). Så det er jo absolutt en måte å gjøre det på, men det er klart, da skal du være litt interessert. Og sånn sett så er jo ikke Power BI noe annet enn presentering av data.»

Sitatet antyder at informanten ikke opplever verktøyet særlig relevant for sin del, men at andre roller i organisasjonen har tatt nytte av det. Samtidig hevder informanten at visualisering av data kan hjelpe til med å besvare spørsmål knyttet til hva man skal gjøre med de store datamengdene:

«På en annen side så kan man jo si at; «okei, du sitter med mye data; hva skal vi bruke de til?». Og da kan jo Power BI være et verktøy som kan hjelpe deg i å visualisere det. (...) Det er jo kanskje en måte å få noe verdi ut av datamengdene på da.»

Informant 3 påstår videre på at kommunikasjonen mellom brukerne og datasjø-teamet ikke har vært helt optimal, og at vedkommende savner mer kontakt med teamet som utvikler rapportene i Power BI:

*«Vi er totalt avhengig av de fra datasjøen for å få ut data (...). Det er veldig lang avstand mellom oss og de som jobber i datasjøen. Jeg vet ikke hva de heter og har aldri snakket med dem. De forstår heller ikke hva slags type data vi har, for når de kommer tilbake med den er den i utgangspunktet bare feil fordi de ikke forstår de ulike parameterne. De har for eksempel trodd at vannmengde var en aggregert mengde, når den faktisk er en type øyeblikksdata (hastighet * diameter = mengde). Så der har vi hatt et problem. Vi vet ikke selv hvordan vi endrer det, så da må vi vente til de har tid og nå har jeg ikke hørt noe fra de på 2 uker. Da går det veldig treg.»*

Sitatene antyder at kommunikasjonen mellom fagpersonenes behov og utviklerne fra datasjøen-teamet ikke har vært helt optimal. Flere av informantene har ikke sett helt behovet for løsningen enda, noe som kan binde kostnader til utvikling og kursing av løsninger som ikke blir tatt i bruk av de antatte brukerne. God kommunikasjon mellom fagpersoner og utviklere virker derfor å være helt sentralt for å skulle lykkes med implementering av nye automatiserte løsninger.

Oppsummering av rapportering

Basert på intervjuene og innledende samtaler, består rapporteringsfunksjonen i VA-etaten av mye manuelt arbeid grunnet systemer som er dårlig integrerte. Etersom manuelt arbeid er ressurskrevende, har VA-etaten ønsker om å automatisere produksjon av rapporter. Anvendelse av historiske data fra datasjøen er en sentral del for å kunne automatisere denne prosessen. Det er nylig utviklet en løsning i verktøyet Power BI, utført av tverrfaglige team fra datasjøen og brukerne i VA-etaten. På tross av dette antyder informantene at de ikke har sett helt nytten av verktøyet enda, noe som kan indikere at kommunikasjon og samhandling ikke har vært helt optimal. Dette virker å være en forutsetning for å utvikle gode automatiserte løsninger.

5.6 Oppsummering av funn

For å oppsummere antyder funnene at Big Data anvendes i nokså stor grad i flere styringsfunksjoner. Sensordata i sanntid fra SCADA-systemet anses å være 'synet' til VA-etaten, og avhengigheten til teknologi er tilsynelatende stor ettersom store deler av infrastrukturen ligger under bakken. Basert på sitatene kan dette 'synet' bidra til bedre overvåking og kontroll ettersom man får bedre oversikt og kan benytte nye verktøy til å identifisere avvik i driftsprosesser. Dette 'synet' ser videre ut til å ha innvirkning på planleggingsfunksjonen, hvor planlegging av tilsyn og vedlikehold endres fra rutinebasert til mer behovsbasert. Planlegging rundt utskiftning av ledninger kan derimot bli betydelig mer kompleks som følge av hensyn til eksterne aktører og et konsekvensorientert fokus. Informantene har store forhåpninger til at data fra eksterne kilder, samt nye analyseverktøy, vil bidra til bedre konsekvensorientert planlegging i fremtiden.

Funn indikerer videre at Big Data spiller en viktig rolle i beslutningstaking. Informasjon om nettet i sanntid, gir beslutningstakere tidsrelevant informasjon som danner et godt grunnlag for å ta gode beslutninger. Likevel anses intuisjon og erfaring å være viktige faktorer i beslutningsprosessen. Informantene hevder samtidig at data må være av høy kvalitet, altså korrekt og nøyaktig, for å kunne ta gode og riktige beslutninger.

Til resultatmåling er måledata i sanntid fra SCADA-systemet avgjørende for å kunne måle KPI'er. Samtidig indikerer funnene at prosessen med å hente ut enkelte KPI'er preges av mye manuelt arbeid ettersom systemer er dårlig integrert. Til slutt antyder funn at funksjonen rapportering også preges av mye manuelt arbeid. VA-etaten har et ønske om å benytte historiske data fra datasjøen til å automatisere produksjon av rapporter, og en løsning er nylig utviklet gjennom verktøyet Power BI. Sitatene indikerer derimot at ikke alle informantene er like fornøyd med løsningen, noe som kan antyde at kommunikasjon og samhandling mellom utviklere og brukere kunne vært bedre.

6. Diskusjon

I denne delen skal vi diskutere implikasjoner knyttet til funnene vi presenterte i kapittel 5. Basert på funnene ønsker vi å trekke ut sentrale innsikter som er relevante for å forstå samspillet mellom styring og Big Data, og hvilke muligheter og utfordringer Big Data gir for organisatorisk styring. Samlet sett indikerer funnene at Big Data kan være et verdifullt fenomen for organisatorisk styring. Denne sammenhengen kommer tydelig frem i funnene hvor Big Data i flere av styringsfunksjonene, særlig overvåking og kontroll, virker å være helt avgjørende for å kunne utøve styring på en god måte. I de neste avsnittene skal vi diskutere dette nærmere.

6.1 Big Data er ‘synet’ som anvendes til styring

Funnene indikerer at data er ‘synet’ til VA-etaten, og uten dette vil etaten være ‘blind’. Bakgrunnen for disse uttalelsene antas å være begrunnet i at infrastrukturen som VA-etaten forvalter, har et stort omfang og består av veldig mange ulike komponenter. Det er blant annet rundt 2200 km med vann- og avløpsledninger, 200 pumpestasjoner og 40 høydebasseng som skal overvåkes for å kontrollere at driften går som den skal. Store deler av ledningsnettets ligger under bakken, noe som gjør det veldig utfordrende å vite noe om tilstanden på nettet uten å benytte data som kan gi målinger og innsikt. Komponenter som er tilgjengelig over bakken, slik som pumpestasjoner og renseanlegg, krever mye ressurser dersom det skulle overvåkes fysisk og uten anvendelse av data.

For styringsfunksjonen overvåking og kontroll ser vi at data - fortrinnsvis fra sensorer - er helt avgjørende for å holde oversikt over VA-nettet. Tilgang på sensordata i sanntid blir beskrevet av informantene som en ‘milepæl’ i bruken av data i VA-etaten, noe som kan antyde at bruken av slike data har medført et radikalt skifte for denne styringsfunksjonen. Intervjudata sier lite om hvordan denne funksjonen ble utført tidligere, men det er mulig å tenke seg at den i større grad var preget av ansattes intuisjon og erfaring om hvordan de ‘antok’ at tilstanden på nettet var. Basert på funn kan det virke som at data har endret denne praksisen stort, hvor man i dag benytter varierte måleverdier fra sensorer, slik som vannmengde, trykk og temperatur. Dette gir mer konkret informasjon om tilstanden på nettet som kan brukes til å holde oversikt og identifisere avvik. Etersom mye av infrastrukturen befinner seg under bakken, ser måledata ut til å gi en bedre oversikt over nettet enn hva man tidligere har klart ved kun intuisjon og

erfaring. Big Data virker derfor i denne konteksten å ha en direkte effekt på overvåking og kontroll ved at det gir bedre oversikt over nettet.

Sensordata i sanntid bidrar til å gi innsikt og tidsrelevant informasjon om hva som skjer på nettet til enhver tid. Denne egenskapen ved Big Data ser ut til å gjøre overvåkingsfunksjonen mer tilpasningsdyktig, noe som gir mulighet til å agere raskere dersom man oppdager driftsstans eller andre kritiske feil. For en kommunal driftsorganisasjon kan dette være veldig verdifullt, ettersom de ønsker å minimere driftsstans på nettet og opprettholde en stabil og robust drift. Denne verdien av tidsrelevant data til styring, virker å være i godt samsvar med tidligere teori (Nielsen, 2018; Rikhardson & Yigitbasioglu, 2018).

I tillegg til interne sensordata, indikerer funnene at VA-etaten ønsker å benytte mer eksterne data for å bedre omfavne eksterne variabler som påvirker driften. Vær og havnivå er eksempler på faktorer som har innvirkning på hvor mye vann som flyter gjennom nettet. Ved å sammenstille slike data med interne måledata, er håpet å få bedre oversikt og forståelse for ulike driftsprosesser. Variasjonen i datamengden kan dermed se ut til å gi et *bredere* 'syn' som i større grad fanger opp og gir oversikt over forskjellige faktorer som har innvirkning på driften. På den måten kan man bedre omfavne og forstå det teori omtaler som 'hele populasjoner' (Rikhardson & Yigitbasioglu, 2018; Kitchin, 2014), altså omfatte alle relevante data istedenfor kun et lite utvalg.

I forlengelse av denne diskusjonen kan man stille spørsmål ved om et bredere 'syn' faktisk vil føre til bedre kontroll. På en side er det naturlig å tenke at mer data gir bedre kontroll ettersom man får en bredere oversikt til å forstå hvilke variabler som påvirker forskjellige driftsprosesser. På en annen side kan dette føre til at man oppdager nye faktorer og nye problemer som man tidligere ikke har tatt hensyn til. Selv om det i utgangspunktet er bra å ta hensyn til flere faktorer, kan det samtidig komplisere prosessen betydelig. Dersom det blir for mye å ta hensyn til kan det være større risiko for at menneskets kapasitet mister kontroll over prosessen. Derfor kreves det gjerne bruk av systemer og avanserte analyseverktøy som har bedre kapasitet enn mennesker til å håndtere all informasjonen. Maskinlæring er et verktøy som har potensiale til dette, men mennesker må likevel forstå hvilke algoritmer modellen skal bruke og hvordan maskinen jobber for å få frem resultater. Dersom man ikke forstår hvordan disse systemene eller verktøyene fungerer, kan man risikere å innbille seg at man har kontroll når man egentlig ikke helt forstår hvordan systemet har generert frem resultatene.

Dersom data skal fungere som 'synet' til VA-etaten, forutsetter det at data er av høy kvalitet - det teorien beskriver som korrekt og troverdig (Chae et al., 2014). Forholder man seg ukritisk til kvaliteten, kan man risikere at det man 'ser' gjennom data ikke stemmer med virkeligheten, noe som kan medføre iverksettelse av feil tiltak. Teori hevder at Big Data har en tendens til å bestå av usikker kvalitet ettersom mye data er ustrukturert og kan være vanskelig å vurdere gyldigheten av (Lee, 2017). Dette medfører derfor en trussel mot troverdigheten til Big Data. På tross av denne trusselen, hevder informantene at datakvaliteten i VA-etaten er god og at de anser den som troverdig. Dette kan skyldes at VA-etaten, gjennom organiseringer som datasjøen, enn så lenge virker å ha god kontroll på hvilke data de bruker og hvor denne kommer fra. Skal etaten derimot fortsette sin ønskede utvikling mot å bli mer datadreven, er det viktig at de fortsetter å sikre høy kvalitet på data for å kunne bruke dette som 'synet' til overvåking og kontroll. En felles datakatalog med god struktur kan være en måte å oppnå dette på.

Videre antyder funn at dette 'synet' også har innvirkning på planleggingsfunksjonen. Funnene indikerer at anvendelse av sensordata i sanntid bidrar til å endre fokuset fra rutinebasert til mer behovsbasert planlegging. Ettersom sensordata gir oversikt og kontroll over komponenter i sanntid, er det i mindre grad nødvendig med fysiske og rutinemessige tilsyn. Ved å overvåke tilstanden til komponentene i et digitalt system, kan man heller fokusere på å planlegge tilsyn der det er behov. Systemet kan gi en oversikt over komponenter som trenger vedlikehold eller utskiftning, og på den måten kan man utføre tiltak etter behov. Dette vil bidra til mer ressurseffektiv drift ettersom ressursene benyttes fullt ut til å fikse driftsproblemer, fremfor å bruke tid på unødvendige og overflødige rutinekontroller.

Samtidig er det andre deler av planleggingsfunksjonen som per i dag anvender Big Data i mindre grad. Funn indikerer at planlegging rundt utskiftning av ledninger er en komplisert prosess som krever hensyn til flere eksterne samfunnsaktører. I tillegg har VA-etaten endret fokus til å tenke på konsekvenser fremfor å kun skifte ledninger med mest brudd. For å ivareta dette fokuset, samt ta hensyn til eksterne aktører, kreves det tilgang til store mengder eksterne data som gir informasjon om planer og implikasjoner for disse aktørene. Funn antyder at VA-etaten ikke har tilgang på slik data i tilstrekkelig grad per i dag, og at dette kan være grunnen til at anvendelsen av Big Data virker å være mindre her. Bakgrunnen til dette kan være at det kreves tilgjengeliggjøring, strukturering og bearbeidelse av eksterne data for å kunne bruke den til analyse. Teori hevder at den store kompleksiteten og ustrukturerte formen til eksterne data kan gjøre det overveldende å jobbe med disse sammenlignet med tradisjonelle databaser

(Appelbaum et al., 2017). Her kan det derimot tenkes at datasjøen til kommunen kan spille en viktig rolle for å realisere tilgjengeliggjøring og bruk av slike data, ettersom de har erfaring med bearbeidelse og strukturering av store datamengder. Informantene virker uansett å håpe på at mer tilgang til Big Data vil kunne bidra til å gjøre bedre prioriteringer i fremtiden.

På en annen side kan det diskuteres om bruk av Big Data faktisk gjør planlegging og prioritering bedre. Økt tilgang til eksterne data medfører også at man får en betydelig mer kompleks helhet å ta hensyn til, som diskutert tidligere. Her kreves det avansert analyse, og tidligere teori hevder at bare det å velge ut hvilke data man skal benytte kan være en stor utfordring (Nielsen, 2018). Etter hvert som man tar hensyn til flere faktorer, vil det muligens også kreves en form for vekting mellom hvilke aktører man skal ta mest hensyn til. Dette gjør det enda mer utfordrende å prioritere og fatte beslutninger i planleggingsprosessen. Når prosessen med å prioritere blir mer omfattende og komplisert, vil ikke lenger menneskets kapasitet til å håndtere de store datamengdene være god nok. Det trengs derfor nye analyseverktøy som har bedre kapasitet til å finne optimale løsninger og prioriteringer enn hva mennesker klarer.

For å kunne anvende Big Data i planlegging på en måte som gir verdi, virker det å være en forutsetning at man benytter teknologi som har kapasitet til å analysere all informasjonen. Maskinlæring kan være en teknologi som bidrar her, men slike verktøy trenger ofte å være lokalt tilpasset og kan kreve store ressurser til utvikling. Teori hevder også at bruk av slike verktøy gjerne krever nye ferdigheter og kompetanse, spesielt innenfor statistisk databehandling (Rikhardson & Yigitbasioglu, 2018). Det kreves også kompetanse på matematisk modellering og algoritmer for å kunne utvikle og benytte maskinlæring på en formålstjenlig måte. Utvikling og utvelgelse av riktige algoritmer til modellering er også en utfordrende jobb, spesielt etter hvert som antall variabler øker. For å hente ut verdier fra slike verktøy må derfor VA-etaten og andre organisasjoner som ønsker å benytte Big Data sikre seg at de har denne kompetansen på plass.

Videre indikerer funn at Big Data har innvirkning på beslutningstaking, ettersom den henger tett sammen med både overvåking og kontroll, og planlegging. Basert på overvåking gjør man beslutninger om hvordan man skal agere, og planlegging handler i stor grad om å prioritere og ta gode beslutninger. 'Synet' som brukes i disse funksjonene ser derfor ut til å være sentralt for å kunne ta gode beslutninger. Funnene indikerer at bruk av Big Data bidrar til å gjøre

funksjonen mer datadrevet, og at data i større grad brukes som beslutningsgrunnlag. Dette virker å være i tråd med det teori omtaler som 'evidensbasert beslutningstaking' (Holsapple et al., 2014).

Teori hevder at en sentral utfordring ved organisasjoner som blir mer datadrevne er at man kan få for stor tillit til dataen, og at viktigheten av menneskelig intuisjon og erfaring kan bli nedvurdert (Quattrone, 2016; Rikhardson & Yigitbasioglu, 2018). Funn indikerer at man ikke kan ta beslutninger basert på data alene, men at man også er avhengig av kunnskap, erfaring og intuisjon, spesielt for å kontrollere datagrunnlaget. Dette er i tråd med litteratur belyst i teorikapittelet (Nielsen, 2018; Quattrone, 2016; Rowley, 2007). Et av funnene belyser hvordan det å ta beslutninger basert på data ukritisk, altså uten å bruke erfaring og intuisjon, kan gi store negative konsekvenser. I VA-etaten ble et renseanlegg bygget betydelig overdimensjonert som følge av ukritisk bruk av data. I forlengelse av slike hendelser kan man derfor stille spørsmål ved om anvendelse av Big Data faktisk fører til bedre beslutninger. Dersom man ukritisk forholder seg til data som beslutningsgrunnlag, kan man risikere at beslutningene tvert imot blir verre. Viktigheten av å bruke menneskelig skjønn, intuisjon og erfaring i beslutningsprosessen kommer derfor tydelig frem i funnene.

Samtidig kan datavegring blant de ansatte være en trussel mot organisasjoner som ønsker å bli mer datadrevet. En sentral utfordring i alle organisasjoner er hvordan man skal få med seg de ansatte på store endringer, spesielt endringer som krever ny teknologisk kompetanse fra de ansatte. Ifølge enkelte informanter er VA-bransjen generelt nokså tradisjonell og litt gammeldags i noen måter å jobbe på, noe som kan bety at overgangen til nye digitale arbeidsmetoder kan være stor. Sunn skepsis til data er sunt, men dersom de ansatte vegrer seg for å benytte digitale verktøy i arbeidet så vil man heller ikke oppnå gevinstene som ny teknologi gir. For å oppnå de ønskede gevinstene ved anvendelse av Big Data og datadrevne verktøy, må derfor organisasjoner ha fokus på god implementering av digitale arbeidsmetoder og arbeidskultur.

Funn indikerer videre at resultatmåling er en funksjon som til en viss grad anvender Big Data. På en side har data en nokså direkte effekt på resultatmåling, ettersom sensordata fra SCADA-systemet bidrar til å hente ut måleverdier som er nødvendige for å måle enkelte KPI'er. Ved måling av for eksempel totalt vannforbruk, er sensormålinger sentralt. På en annen side antyder funn at flere KPI'er krever sammenstilling av data, noe som i stor grad utføres manuelt

ifølge informantene. For å få til en automatisering av denne prosessen trengs det blant annet nye systemer som kan sammenstille data på en bedre måte.

Etter hvert som datamengdene øker, kan det imidlertid stilles spørsmål ved om dette kommer til å utfordre dagens resultatmål. Data fra stadig flere kilder kan kanskje føre til at man får ny informasjon som utfordrer tradisjonelle måter å tenke på. Dette kan igjen ha innvirkning på hva som egentlig er relevant å måle lenger. Kanskje blir det ikke lenger relevant å måle for eksempel antall prosent nedgang i fremmedvann, dersom man ved hjelp av data finner ut at det er andre faktorer som har større innvirkning på prosessen. Selv om det er vanskelig å vite hva som er relevant å måle i fremtiden, virker Big Data å ha potensiale til å endre dette.

Til slutt indikerer funn at rapporteringsfunksjonen er i startfasen med å anvende mer Big Data for å automatisere rapporter. Per nå er funksjonen derimot preget av en del manuelt arbeid. Ettersom manuelt arbeid er utsatt for menneskelige feil kan dette være en potensiell feilkilde i disse oppgavene. Dette kan for eksempel skje ved at man leser av feil tall under innhenting og sammenstilling av data fra forskjellige systemer. VA-etaten har hatt et ønske om å automatisere produksjon av rapporter, og har utviklet noen automatiserte løsninger i Power BI. Automatiserte datasystemer er ikke underlagt de samme begrensningene som mennesker og er derfor langt mer nøyaktige. Økt bruk av Big Data og automatisering kan dermed bidra til å redusere disse feilkildene og samtidig gjøre arbeidet mer effektivt.

Samtidig kan overgangen til automatiserte løsninger også medføre nye utfordringer og feilkilder. Blant annet indikerer funn at det kreves godt samspill mellom utviklere og fagfolk ved utvikling av slike løsninger, noe som er viktig for at løsningene skal oppleves nyttige av brukerne. Intervjuene belyste blant annet eksempler på at enkelte informanter ikke stolte helt på at de som utviklet rapportene brukte riktige parameter for formålet. Dette kan sees på både som en organisatorisk utfordring, samt en teknisk utfordring. En av årsakene til denne skepsisen kan være mangel på kommunikasjon blant fagfolk og utviklere. Det kan tenkes at i en fagorganisasjon som VA-etaten er det mye erfaringsbasert kunnskap, noe som kan gjøre at dette samarbeidet blir desto viktigere. Dersom ikke dette samspillet er optimalt, kan man risikere å binde kostnader til utvikling og opplæring av løsninger som ikke oppfattes som nyttige av brukerne.

6.2 Big Data påvirker ulike deler av styring og på ulike tidspunkt

Funn indikerer at Big Data kan gi store muligheter for flere av styringsfunksjonene i VA-etaten. Samtidig ser det ut til at enkelte funksjoner per nå har kommet lengre med anvendelse av Big Data enn andre. Det kan være mange grunner til hvorfor, og i de neste avsnittene skal vi prøve å diskutere noen mulige årsaker til dette.

En organisatorisk årsak til at noen styringsfunksjoner kan ha kommet lenger med bruk av Big Data, kan skyldes fokuset til organisasjonen. Det overordnede fokuset til en kommunal driftsorganisasjon handler gjerne om å minimere driftsstans og driftsforstyrrelser. VA-etaten skal levere stabile og robuste tjenester for abonnenter, samtidig som de tar hensyn til andre offentlige aktører med tanke på utbyggingsprosjekter og andre hendelser. De viktigste styringsfunksjonene virker derfor å være knyttet til overvåking, planlegging og beslutning av tiltak som minimerer driftsforstyrrelser. Dette kan være noe av grunnen til at funksjonen overvåking og kontroll, samt noen deler av planlegging og beslutningstaking, kan å ha kommet lengre med anvendelse av Big Data. Rapportering og resultatmåling anses likevel å være sentrale styringsfunksjoner, men kanskje ikke like driftskritiske som de andre funksjonene i denne konteksten.

Forskjeller i kompleksiteten til funksjonene kan være en annen potensiell årsak til forskjeller blant funksjonene. Overvåking og kontroll handler mye om å ha oversikt og identifisere avvik i driftsprosesser. Når man har tilgang på sensordata som gir konkrete måleverdier, kan det bli betydelig enklere å holde oversikt og kontroll på hva som skjer langs nettet til enhver tid. Big Data virker derfor å ha en nokså direkte innvirkning på funksjonen, spesielt i denne konteksten hvor store deler av infrastrukturen ligger under bakken. Et bedre 'syn' i form av sensordata i sanntid, kan derfor se ut til å gi nokså raske gevinster for overvåking og kontroll i en slik organisasjon hvor det er relativt stor avstand mellom infrastrukturen og menneskene.

Når man derimot i neste fase skal planlegge og ta beslutninger knyttet til hvilke tiltak man skal iverksette, tyder funn på at denne prosessen er mer omfattende og krever mer avansert analyse. Etter hvert som mer ekstern data anvendes, kan prosessen med å prioritere bli enda mer komplisert. For å anvende Big Data i denne prosessen på en verdifull måte, ser det ut til å være en forutsetning at man benytter nye systemer og avanserte analyseverktøy. Et eksempel er prosessen med utskifning av ledninger på nettet. Per i dag brukes det få avanserte verktøy for

denne prosessen, men ifølge enkelte informanter er dette noe som er under utvikling. Dette kan være noe av grunnen til at funksjonen planlegging ikke enda har klart å anvende Big Data i like stor grad som overvåking og kontroll. Hvor overvåking og kontroll ser ut til å gi relativt raske gevinster, kan det virke som at planlegging og beslutningstaking er mer komplekse funksjoner som i større grad krever nye tekniske og organisatoriske forutsetninger.

Basert på konteksten i dette caset kan *tid* være en sentral dimensjon for å forstå anvendelse av Big Data. Funn antyder at Big Data i tidlig fase gir en nokså umiddelbar gevinst for overvåking og kontroll, ettersom 'synet' hjelper å få bedre oversikt over infrastrukturen under bakken. Det samme gjelder resultatmåling, hvor data umiddelbart gir en gevinst i form av konkrete målinger. Enkelte planleggingsoppgaver kan også å få en nokså rask gevinst i tidlige faser ettersom 'synet' bidrar til å gjøre tilsyn og vedlikehold mer behovsbasert.

På en annen side kan det se ut som at funksjoner som krever hensyn til mange eksterne faktorer, slik som enkelte deler av planlegging og beslutningstaking, har en kompleksitet ved seg som gjør at det kreves at man får på plass nye avanserte verktøy og mer omfattende strukturer før man kan hente ut verdi. Av den grunn kan det se ut som at det tar litt mer tid før denne funksjonen klarer å anvende Big Data fullt ut. Beslutningstaking kan sees i sammenheng med planlegging og er derav også en relativt kompleks funksjon. Selv om 'synet' kan antydes å ha gitt et bedre informasjonsgrunnlag for å ta gode beslutninger, må data gjerne kvalitetssikres opp mot intuisjon og erfaring før det kan fattes endelige beslutninger. Derfor kan det tenkes at det også tar noe mer tid før denne funksjonen klarer å benytte Big Data fullt ut.

Oppsummert kan det virke som at Big Data påvirker ulike deler av styring og på ulike tidspunkt. Funn indikerer at ikke alle styringsfunksjoner klarer å anvende Big Data like raskt, ettersom kompleksiteten i enkelte funksjoner ser ut til å kreve noen nye organisatoriske og tekniske strukturer. Dette er funn som tidligere teori om Big Data og styring ikke har belyst, og er derfor noe vi hevder kan være et bidrag til forskning på fenomenet.

7. Konklusjon

Formålet med denne masterutredningen har vært å undersøke hvordan Big Data anvendes i organisatorisk styring. Fremveksten av nye digitale enheter, medfører at det produseres mer data nå enn noen gang tidligere. Gitt at Big Data kan anses som en kilde til informasjon, er dette et fenomen som potensielt kan ha store innvirkninger på organisatorisk styring. På tross av dette, er det lite empiriske studier som har studert sammenhengen mellom Big Data og styring. Vi har derfor ønsket å belyse følgende problemstilling: *“Hvordan anvendes Big Data i organisatorisk styring?”*

Våre funn indikerer at Big Data kan anses som ‘synet’ til VA-etaten, noe som virker å ha stor innvirkning på styring i organisasjonen. Gjennom anvendelse av sensordata i sanntid får etaten bedre oversikt over ulike komponenter på nettet, noe som kan anses å være veldig verdifullt i denne konteksten hvor mye av infrastrukturen ligger under bakken. Dette ser ut til å gi en umiddelbar gevinst til styringsfunksjonen overvåking og kontroll ved at man får bedre oversikt. Funnene indikerer videre at Big Data har innvirkning på andre styringsfunksjoner, blant annet gjennom mer presise resultatmålinger, et bredere beslutningsgrunnlag, og muliggjøring for mer behovsbasert planlegging av tilsyn og vedlikehold.

På en annen side kan det argumenteres for at økt bruk av Big Data øker kompleksiteten i flere av funksjonene. Hensyn til stadig flere faktorer gjør at planlegging og beslutningstaking blir mer komplisert, noe som kan gjøre det mer utfordrende å prioritere og fatte beslutninger. For å håndtere komplekse sammensetninger av data, virker det derfor å være en forutsetning at man klarer å benytte teknologi og avanserte analyseverktøy som kan finne optimale løsninger. Dette krever gjerne at man legger innsats og ressurser i å utvikle konkrete og lokalt tilpassede løsninger. Det er også viktig å sikre at man har riktig kompetanse i form av statistisk databehandling og modellering for å kunne jobbe med slike verktøy. Basert på dette kan det virke som at mer komplekse styringsfunksjoner trenger noe lengre tid før det klarer å realisere gevinster knyttet til anvendelse av Big Data.

Basert på funn fra denne studien, kan det se ut som at tid er en viktig dimensjon for å forstå forskjeller i anvendelse av Big Data i de ulike styringsfunksjonene.

7.1 Implikasjoner og videre forskning

I denne studien vil vi hevde å ha gjort flere interessante funn, men anser særlig to momenter å kunne være bidrag til forskning innen Big Data og styring. Vår studie antyder at Big Data har innvirkning på styring, men (i) ulike deler av styring og (ii) på ulike tidspunkt. Big Data virker å gi en umiddelbar gevinst til overvåking og kontroll ved at man får bedre oversikt, særlig i en bransje hvor avstanden mellom mennesker og kapital er stor. Samtidig medfører økt bruk av eksterne data at enkelte styringsfunksjoner, slik som planlegging og beslutningstaking, kan bli betydelig mer komplekse. Dette krever ny kompetanse og bruk av avanserte analyseverktøy, noe som kan føre til at det tar noe lengre tid før man kan realisere gevinster fra bruk av Big Data i disse funksjonene.

Ettersom vi var interessert i å søke etter dybdekunnskap om hvordan Big Data anvendes i styring, valgte vi å gjennomføre en casestudie av VA-etaten i Bergen kommune. Dette bidro til å belyse detaljer og nyanser om fenomenet i en gitt kontekst, noe som var viktig for å forstå *hvordan* Big Data anvendes i organisatorisk styring. En naturlig begrensning av en casestudie er naturligvis at den er utført i en enkelt kontekst, og dermed ikke er direkte overførbart til andre kontekster. Funnene kan dermed ikke generaliseres, og det er usikkert om de samme funnene hadde vært gjeldende i andre organisasjoner. Vi anser derimot at funnene kan være gjeldende for liknende organisasjoner, spesielt andre virksomheter i VA-bransjen eller andre bransjer hvor avstanden mellom mennesker og kapital er stor.

I videre forskning på fenomenet kunne det vært nyttig å gjennomføre en mer langsiktig *longitudinell* studie av en lignende organisasjon for å se på hvordan utviklingen og adaptasjonen av Big Data utartes over tid. Ettersom vi i denne studien argumenterer for at tid er en viktig dimensjon for å forstå anvendelsen av Big Data, kan dette være særlig interessant. Det kan potensielt avdekke nye muligheter og utfordringer for anvendelse av Big Data til styring. I vår oppgave har vi fokusert på bruk av data i konteksten av VA-bransjen i offentlig sektor. Det kunne derfor vært interessant å gjennomføre en liknende studie i en privat virksomhet eller i en annen bransje. Dette kunne potensielt gi nye innsikter i fenomenet. Videre har vi i denne oppgaven også fokusert utelukkende på operasjonelle styringsfunksjoner. Det kunne derfor vært interessant og gjennomført en studie som fokuserte på finansielle styringsfunksjoner, gjerne i en offentlig virksomhet hvor kostnadsnivået er sentralt for å sikre ressurseffektiv drift.

8. Litteraturliste

- Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal Of Accounting Information Systems*, 25, 29-44. <https://doi.org/10.1016/j.accinf.2017.03.003>
- Atkinson, A., Kaplan, R., & Young, M. (2011). *Management accounting*. New Jersey: Prentice Hall
- Bergen kommune (2020, 9. juni). *Rent vann til folk og fjord*. Hentet 20. mai 2021 fra: <https://www.bergen.kommune.no/omkommunen/avdelinger/vannog-avlopsetaten/om-oss/rent-vann-til-folk-og-fjord>
- Bergen kommune (u.d.-a). *Handlings- og økonomiplan 2021-2024*. Hentet 20. mai 2021 fra: https://pub.framsikt.net/2021/bergen/bm-2021-_hop_kortversjon_21-24/#/
- Bergen kommune (u.d.-b). *Digitaliseringsstrategi Bergen kommune 2021-2025 (høringsutkast)*. Hentet 20. mai 2021 fra: <https://www.bergen.kommune.no/politikere-utvalg/api/fil/3579832/Digitaliseringsstrategi-Bergen-kommune-2021-2025->
- Bergen kommune (u.d.-c) *Hovedplan for avløp og vannmiljø 2019–2028*. Hentet 20. mai 2021 fra: <https://www.bergen.kommune.no/publisering/api/filer/T542342237>
- Bergen kommune (u.d.-d). *Bergen kommune, SDI og WeB*. Upublisert manuskript. SDI.
- Bergen Vann (u.d.). *Organisasjon*. Hentet 20. mai 2021 fra: <https://www.bergenvann.com/ansatte/organisasjon/>
- Berry, A. J., & Otley, D. T. (2004). Case-based research in accounting. *The real life guide to accounting research*, 231-255
- Bhimani, A., & Willcocks, L. (2014). Digitisation, 'Big Data' and the transformation of accounting information. *Accounting And Business Research*, 44(4), 469-490. <https://doi.org/10.1080/00014788.2014.910051>
- Bouvet (2019, 24. mars). *Et datasjøprosjekt til inspirasjon*. Hentet 20. mai 2021 fra: <https://www.bouvet.no/bouvet-deler/et-ordentlig-datasjøprosjekt>
- Caglio, A. (2003). Enterprise Resource Planning systems and accountants: towards hybridization?. *European Accounting Review*, 12(1), 123-153. <https://doi.org/10.1080/0963818031000087853>
- Chae, B., Yang, C., Olson, D., & Sheu, C. (2014). The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. *Decision Support Systems*, 59, 119-126. <https://doi.org/10.1016/j.dss.2013.10.012>

-
- Constantinou, I.D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, 30(1), 44-57. <https://doi.org/10.1057/jit.2014.17>
- Davenport, T. H., Bart, P., Bean, R. (2012). How 'Big Data' is Different. *MIT Sloan Management Review*, 54(1)
- Davenport, T.H. & Kim, J. (2013), *Keeping up with the Quants: Your Guide to Understanding and Using Analytics*, Boston: Harvard Business Review Press
- Eilars, S. (2013, 12. september). *Beyond Volume, Variety and Velocity is the Issue of Big Data Veracity*. Inside Big Data. Hentet 20. mai 2021 fra: <https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management*, 35(2), 137-144
- Granlund, M. (2011). Extending AIS research to management accounting and control issues: A research note. *International Journal Of Accounting Information Systems*, 12(1), 3-19. <https://doi.org/10.1016/j.accinf.2010.11.001>
- Grønmo, S. (2021). Tverrsnittstudie. I Store norske leksikon. Hentet 20. mai fra: <https://snl.no/tverrsnittstudie>
- Grover, V., Chiang, R. H., Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, 35(2), 388-423. <https://doi.org/10.1080/07421222.2018.1451951>
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, California: Sage Publications
- Holmes, D. E. (2017). *Big data: a very short introduction*. Oxford: Oxford University Press.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*, 64, 130-141. <https://doi.org/10.1016/j.dss.2014.05.013>
- IMA. (2008). *Statements on Management Accounting Definition of Management Accounting*. Hentet 20. mai 2021 fra: <https://www.imanet.org/-/media/6c984e4d7c854c2fb40b96bfbe991884.ashx?as=1&mh#:~:text=Management%20accounting%20is%20a%20profession,tation%20of%20an%20organization's%20strategy.>
- Johnson, H. and Kaplan, R. (1987). *Relevance Lost: The Rise and Fall of Management Accounting*. Boston: Harvard Business School Press.
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), <https://doi.org/10.1177/2053951714528481>

-
- Knudsen, D. (2020). Elusive boundaries, power relations, and knowledge production: A systematic review of the literature on digitalization in accounting. *International Journal Of Accounting Information Systems*, 36, 100441. <https://doi.org/10.1016/j.accinf.2019.100441>
- Lambert, F. (2020, 22. april). *Tesla drops a bunch of new Autopilot data, 3 billion miles and more*. Hentet 20. mai 2021 fra: <https://electrek.co/2020/04/22/tesla-autopilot-data-3-billion-miles/>
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70), 1.
- Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3), 293-303. <https://doi.org/10.1016/j.bushor.2017.01.004>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park: Sage Publications
- Marr, B. (2018, 21. mai). *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read*. Hentet 20. mai 2021 fra: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/>
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, 90(10), 60-68.
- Mekki, A. (2020). *Datasjø Bergen 13.000 sensorer for Vann/avløp*. Hentet 20. mai 2021 fra: <https://geoforum.no/wp-content/uploads/2020/01/Datasj%C3%B8.-Bergen-kommune.pdf>
- Microsoft (u.d.). *Hva er maskinlæring?*. Hentet 20. mai 2021 fra: <https://azure.microsoft.com/nb-no/overview/what-is-machine-learning-platform/>
- Moll, J., & Yigitbasioglu, O. (2019). The role of internet-related technologies in shaping the work of accountants: New directions for accounting research. *The British Accounting Review*, 51(6), 100833. <https://doi.org/10.1016/j.bar.2019.04.002>
- Nielsen, S. (2018). Reflections on the applicability of business analytics for management accounting – and future perspectives for the accountant. *Journal Of Accounting & Organizational Change*, 14(2), 167-187. <https://doi.org/10.1108/JAOC-11-2014-0056>
- Otley, D. (2008). Did Kaplan and Johnson get it right?. *Accounting, Auditing & Accountability Journal*, 21(2), 229-239. <https://doi.org/10.1108/09513570810854419>
- Prasad, A., & Green, P. (2015). *Organizational Competencies and Dynamic Accounting Information System Capability: Impact on AIS Processes and Firm Performance*, 29(3), 123-149. <https://doi.org/10.2308/isys-51127>
- Qu, S., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research In Accounting & Management*, 8(3), 238-264. <https://doi.org/10.1108/11766091111162070>

-
- Quattrone, P. (2016). Management accounting goes digital: Will the move make it wiser?. *Management Accounting Research*, 31, 118-122. <https://doi.org/10.1016/j.mar.2016.01.003>
- Redman, Thomas C., (2013). *Data Driven: Profiting From Your Most Important Business Asset*. Boston: Harvard Business Press
- Reinsel, D., Gantz, J., Rydning, J. (2017, april). *Data Age 2025: The Evolution of Data to Life-Critical. Don't focus on Big Data; Focus on the Data That's Big*. Hentet 20. mai 2021 fra: <https://www.import.io/wp-content/uploads/2017/04/Seagate-WP-DataAge2025-March-2017.pdf>
- Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal Of Accounting Information Systems*, 29, 37-58. <https://doi.org/10.1016/j.accinf.2018.03.001>
- Rom, A., & Rohde, C. (2007). Management accounting and integrated information systems: A literature review. *International Journal Of Accounting Information Systems*, 8(1), 40-68. <https://doi.org/10.1016/j.accinf.2006.12.003>
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal Of Information Science*, 33(2), 163-180. <https://doi.org/10.1177/0165551506070706>
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students*. Harlow: Pearson Education.
- Scapens, R. W. (2004). Doing case study research. *The real life guide to accounting research*, 257-279.
- Simon, P. (2013) *Too Big to Ignore: The Business Case for Big Data*. New Jersey: Wiley.
- SSB (u.d.). *Kommunefakta Bergen*. Hentet fra: <https://www.ssb.no/kommunefakta/bergen>
- The Economist (2017, 6. mai). *The world's most valuable resource is no longer oil, but data*. Hentet 20. mai 2021 fra: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- Tidemann, A. & Elster, A.C. (2019, 7. juni), Maskinl ring. I *Store norske leksikon*. Hentet 20. mai 2021 fra: <https://snl.no/maskinl%C3%A6ring>
- Warren Jr., J.D., Moffitt, K.C. and Byrnes, P. (2015). *How Big Data will Change Accounting*. *Accounting Horizons*, 29, 397-407. <https://doi.org/10.2308/acch-51069>
- Windeck, D., Weber, J., & Strauss, E. (2013). Enrolling managers to accept the business partner: the role of boundary objects. *Journal Of Management & Governance*, 19(3), 617-653. <https://doi.org/10.1007/s10997-013-9277-2>
- Zhang, J., Yang, X., & Appelbaum, D. (2015). Toward Effective Big Data Analysis in Continuous Auditing. *Accounting Horizons*, 29(2), 469-476. <https://doi.org/10.2308/acch-51070>

Vedlegg 1: Semistrukturert intervjuguide

Innledende spørsmål:

- Hva er din stilling/rolle?
 - Hvilken VA-avdeling tilhører du?
 - Hva er dine mest sentrale arbeidsoppgaver?
- Hvor lenge har du jobbet i VA-etaten / Bergen Vann?

Big Data, dataverktøy og styringsfunksjoner:

- Hvordan bruker du data og dataverktøy til overvåking og kontroll?
- Hvordan bruker du data og dataverktøy til planlegging?
- Hvordan bruker du data og dataverktøy til rapportering?
- Hvordan bruker du data og dataverktøy til resultatmåling?
- Hva opplever du er viktig å tenke på når man jobber med data og datadrevne verktøy?
- Hvilken rolle spiller data og dataverktøy når du skal ta beslutninger?
 - Hvor viktig er det med intuisjon og erfaring når det tas beslutninger?

Oppfølgingsspørsmål:

- Har du eksempel på en situasjon hvor dette er aktuelt?
 - Hva er bra med data og aktuelle verktøy her?
 - Hva er utfordrende med data og aktuelle verktøy her?
 - Hvordan ble dette gjort før?

Vedlegg 2: Informasjonsskriv



Invitasjon til deltakelse i forskningsprosjekt

Som en del av mastergraden ved Norges Handelshøyskole (NHH) er vi to studenter som skriver masteroppgave om; «*Hvordan anvendes Big Data i organisatorisk styring?*» - *En casestudie av VA-etaten i Bergen kommune.*

Formålet med forskningsprosjektet er å undersøke hvordan en organisasjon anvender Big Data og dataverktøy til styring. Vår oppfattelse er at VA-etaten er en spennende organisasjon for å studere dette.

Som en sentral del av prosjektet ønsker vi å gjennomføre intervjuer med personer i Bergen Vann og VA-etaten som bruker Big Data og dataverktøy i sin jobb. Vi er interessert i å høre om hvordan du bruker de teknologiske verktøyene inn mot styringsfunksjoner, og hvilke erfaringer og tanker du har rundt dette. Vi håper at forskningsprosjektet kan bidra med relevant innsikt i dette temaet, både for VA-etaten og for oss på NHH.

Lengden på intervjuet er estimert til 30-45 min.

Vi ønsker å benytte lydopptak av intervjuet for å sikre at vi fanger opp informasjon på en riktig måte. All behandling av informasjonen du deler i intervjuet vil være anonym. Lydopptak og transkriberinger av intervjuene vil kun være tilgjengelig for undertegnede, og slettes etter at forskningsopplegget er avsluttet 01.06.2021.

Med vennlig hilsen,

Paul Magne Granberg og Henrik Hjelset

Vedlegg 3: Samtykkeskjema

Vil du delta i forskningsprosjektet ”Big Data i styringsfunksjonen”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke hvordan en organisasjon anvender Big Data og data drevne-verktøy i ulike styringsfunksjoner. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

Forskningsprosjektet skal utforske følgende problemstilling;

«Hvordan anvendes Big Data i organisatorisk styring?»

Denne problemstillingen ønsker vi å utforske gjennom å utføre semistrukturerte individuelle dybdeintervju og inviterer deg herved til å delta. Spørsmålene i intervjuet vil være relativt åpne og handle om hvordan en organisasjon anvender Big Data og data drevne-verktøy i ulike styringsfunksjoner. Med styringsfunksjoner mener vi blant annet planlegging, kontrollering, overvåking, måling av organisasjonens ytelse og beslutningstaking.

Om du ønsker å delta i prosjektet ber vi deg om å sende en skriftlig bekreftelse på e-post til henrik.hjelset@student.nhh.no

Hvem er ansvarlig for forskningsprosjektet?

Norges Handelshøyskole (NHH) er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Utvalget er basert på bekvemmelighetsutvalg. Våre kontakter i VA-etaten har anbefalt deg som en aktuell intervjukandidat basert på din erfaring relatert til tema. Vi tar sikte på å intervju ca 8-10 personer.

Hva innebærer det for deg å delta

Hvis du velger å delta i prosjektet innebærer det å være med på et intervju som er estimert til 45 minutter, og vil foregå via videosamtale. Det vil gjøres lydopptak av intervjuet. Dette vil bli slettet når det er transkribert til skriftlig form, og all bruk av informasjon du deler vil være anonymisert. I intervjuet vil det blant annet bli spurt om hvilken stilling du har, og hvordan big data og data drevne-verktøy brukes i arbeidet ditt.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Det er kun studentene som ansvarlig for prosjektet som vil ha tilgang til opplysningene, som vil lagres på en kryptert skytjeneste med flere sikkerhetstiltak for å hindre uautorisert adgang. Så snart intervjuet er transkribert vil alle opplysninger bli anonymisert slik at du ikke vil kunne gjenkjennes.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er rundt 22. juni 2021.

Deltakelse i intervjuet er frivillig, og du kan når som helst trekke ditt samtykke. Du kan ta kontakt med oss dersom du ønsker å trekke ditt samtykke, ha innsyn i datamaterialet fra ditt intervju eller har andre spørsmål.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Norges Handelshøyskole (NHH) har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Norges Handelshøyskole ved Andreas Ulfstein. E-post: Andreas.ulfstein@nhh.no, tlf:99111822
- Vårt personvernombud: personvernombud@nhh.no

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller påtelefon: 55 58 21 17.

Med vennlig hilsen

Andreas Ulfstein
(Veileder)

Paul Magne Granberg
(Student)

Henrik Stokken Hjelset
(Student)

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Big data i styringsfunksjonen*, og har fåttanledning til å stille spørsmål. Jeg samtykker til:

å delta i dybdeintervju

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er

avsluttet (Signert av prosjektdeltaker, dato)

Vedlegg 4: NSD sin vurdering av personvern i prosjektet

Prosjekttittel

Masteroppgave: Anvendelse av Big Data i organisatorisk styring - En casestudie av Vann- og avløpsetaten i Bergen kommune

Referansenummer

204417

Registrert

10.02.2021 av Paul Magne Granberg - Paul.Granberg@student.nhh.no

Behandlingsansvarlig institusjon

Norges Handelshøyskole / Institutt for regnskap, revisjon og rettsvitenskap

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Andreas Ulfsten, andreas.ulfsten@nhh.no, tlf: 55959362

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Paul Magne Granberg, paul.granberg@student.nhh.no, tlf: 47809638

Prosjektperiode

12.03.2021 - 21.06.2021

Status

17.02.2021 - Vurdert

Vurdering (1)

17.02.2021 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 17.02.2021, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte. DEL PROSJEKTET MED PROSJEKTANSVARLIG Det er obligatorisk for studenter å dele meldeskjemaet med prosjektansvarlig (veileder). Det gjøres ved å trykke på "Del prosjekt" i øvre venstre hjørne av meldeskjemaet. MELD VESENTLIGE ENDRINGER Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde: <https://www.nsd.no/personverntjenester/fulle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema> Du må vente på svar fra NSD før

endringen gjennomføres. TYPE OPPLYSNINGER OG VARIGHET Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 21.06.2021. LOVLIG GRUNNLAG Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a. PERSONVERNPRINSIPPER NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), og dataportabilitet (art. 20). NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13. Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned. FØLG DIN INSTITUSJONS RETNINGSLINJER NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32). Ved bruk av databehandler (spørreskjemaleverandør, skylagring eller videosamtale) må behandlingen oppfylle kravene til bruk av databehandler, jf. art 28 og 29. For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon. OPPFØLGING AV PROSJEKTET NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet. Lykke til med prosjektet! Tlf. Personverntjenester: 55 58 21 17 (tast 1)