

Beate Drønnen  
John-Kevin Sydow Smådal  
Thomas Lande

# Utforming av et rammeverk for utvikling av data-drevne makroanalyser i virksomheter

Bacheloroppgave i Digital forretningsutvikling

Veileder: Thomas Østerlie

Mai 2023



Beate Drønnen  
John-Kevin Sydow Smådal  
Thomas Lande

# **Utforming av et rammeverk for utvikling av data-drevne makroanalyser i virksomheter**

Bacheloroppgave i Digital forretningsutvikling  
Veileder: Thomas Østerlie  
Mai 2023

Norges teknisk-naturvitenskapelige universitet  
Fakultet for informasjonsteknologi og elektroteknikk  
Institutt for datateknologi og informatikk



Kunnskap for en bedre verden



# Forord

Denne bacheloroppgaven er skrevet ved Institutt for Datateknologi og Informatikk, ved Norges teknisk-naturvitenskapelige universitet. Oppgaven markerer slutten på et treårig studieløp innen Digital forretningsutvikling, og har blitt skrevet vårsemesteret 2023. Oppgaven er formet ut fra et dypdykk i interessante temaer som er relevant for studieprogrammet.

Vi fikk muligheten til å skrive denne oppgaven i samarbeid med twoday Avento. Etter flere møter med veileder og bedrift ble det klart at vi ønsket å undersøke hvordan virksomheter kunne benytte seg av data-drevne makroanalyser for å møte den uoversiktige situasjonen i verdensmarkedet i dag. Data-drevne makroanalyser har i kjølvannet av covid-pandemien satt preg på, og kommer garantert til å videre være en stor del av virksomheters beslutningsgrunnlag i fremtiden.

Vi ønsker å takke vår kontaktperson Christoffer Haukås i twoday Avento for givende og gode samtaler og et innblikk i virksomhetens drift. Vi ønsker også å rette en stor takk til ansatte i twoday Avento som ønsket å stille til intervju i forskningsprosjektet vårt.

Til slutt ønsker vi å takke vår veileder Thomas Østerlie for hjelp og god oppfølging gjennom hele utformingen av oppgaven.

Dette bachelorløpet har vært en enormt verdsatt og svært lærerik tid. Etter bacheloroppgaven sitter vi igjen med ny verdifull kunnskap og perspektiv. Det har også vært særs givende å utforske et lite omtalt felt i vårt fagområde, som uten tvil kommer til å bli svært aktuelt i fremtiden.

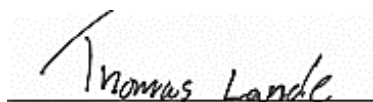
Trondheim, mai 2023



Beate Drønnen



John-Kevin Sydow Smådal



Thomas Lande

# Sammendrag

I et mer uoversiktlig verdensbilde har virksomheter i dag blitt mer avhengig av å ta beslutninger fortløpende, basert på data som øker i mengde og kompleksitet. Mange virksomheter opplever at verdenssituasjonen skaper stor uforutsigbarhet. De har bare de siste årene opplevd utfordringer knyttet til covid-pandemien, krig i Europa, drastiske prisøkninger, klimautfordringer, det grønne skiftet, med mer. Hvordan kan digitalisering hjelpe organisasjoner med disse utfordringene? Og hvordan kan digitalisering bidra til økte resultater i virksomheter?

I denne rapporten skal vi se på mulighetene knyttet til data-drevne makroanalyser og hvordan dette kan bidra til å gi virksomheter et bedre beslutningsgrunnlag. For å skrive oppgaven har vi tatt i bruk Design Science Research Methodology, og har gjennom semistrukturerte intervju og litteratursøk dannet et godt grunnlag for synspunkter og refleksjoner rundt tematikken. Videre har vi synliggjort hva som kreves teknisk og organisatorisk av en virksomhet, og deretter skissert et forslag til et rammeverk for utvikling av data-drevne makroanalyser i virksomheter.

Oppgaven konkluderer med at et rammeverk for utvikling av data-drevne makroanalyser kan være en verdifull ressurs for virksomheter som ønsker å utnytte potensialet i data de alt innehar, eller er i stand til å innhente. Ved å følge rammeverket kan virksomheter oppnå bedre innsikt i omgivelsene og ta mer informerte beslutninger. Studien peker også på mulige fremtidige forskningsområder og videreutvikling av rammeverket.

# Abstract

In an increasingly uncertain global landscape, businesses today have become more reliant on making continuous decisions based on growing volumes and complexity of data. Several organizations are experiencing significant unpredictability due to the world situation. In recent years alone, they have faced challenges related to the COVID-19 pandemic, warfare in Europe, drastic price increases, climate change issues, the green shift, and more. How can digitalization help organizations address these challenges? And how can digitalization contribute to increased performance in businesses?

This report examines the possibilities associated with data-driven macroanalytics and how they can provide businesses with a better decision-making foundation. To write this thesis, we have utilized the Design Science Research Methodology, forming a solid basis for perspectives and reflections through semi-structured interviews and review of literature. Furthermore, we have identified the technical and organizational requirements within a business and outlined a proposal for a framework for developing data-driven macro analyses in organizations.

The thesis concludes that a framework for developing data-driven macro analyses can be a valuable resource for businesses seeking to leverage the potential of the data they already possess or are capable of obtaining. By following the framework, organizations can gain better insights into their surroundings and make more informed decisions. The study also highlights potential future research areas and further development of the framework.

# Innhold

|  |     |
|--|-----|
| Forord .....   | v   |
| Sammendrag .....   | vi  |
| Abstract .....   | vii |
| Figurer .....  | x   |
| Tabeller .....   | x   |
| Forkortelser .....   | xi  |
| 1 Innledning .....   | 12  |
| 1.1 Bakgrunn .....   | 12  |
| 1.2 Problemstilling og forretningsproblem .....                    | 13  |
| 1.3 Avgrensninger .....  | 14  |
| 1.4 Oppgavens oppbygging .....                                     | 14  |
| 1.5 Om oppgavestiller .....  | 15  |
| 2 Teoretisk grunnlag .....   | 16  |
| 2.1 Datavitenskap.....   | 16  |
| 2.2 Data .....   | 16  |
| 2.3 Stordata.....  | 18  |
| 2.4 Dataanalyse .....  | 18  |
| 2.5 Dataanalyseteam .....  | 19  |
| 2.6 CRISP-DM .....   | 21  |
| 2.7 Kunstig intelligens .....                                      | 23  |
| 2.8 Data-drevne metoder .....                                      | 23  |
| 2.9 Virksomheters bruk av data.....                                | 24  |
| 2.10 Benefit Dependency Networking .....                           | 26  |
| 2.11 Strategisk analyse av omgivelser .....                        | 26  |
| 2.12 Makroanalyser .....   | 27  |
| 2.13 Data-drevne makroanalyser .....                               | 28  |
| 2.14 Dataarkitektur .....  | 29  |
| 2.15 Datavarehus og datasjø .....                                  | 31  |
| 3 Metode .....   | 32  |
| 3.1 Forskningstilnærming: Design Science Research Methodology..... | 32  |
| 3.2 Steg 1 Problemidentifisering og motivasjon .....               | 33  |
| 3.3 Steg 2 Formulere mål for og krav til løsningen .....           | 35  |
| 3.3.1 Kravhåndtering .....   | 35  |
| 3.3.2 Interessentanalyse .....                                     | 37  |
| 3.3.3 Kravinnhenting .....   | 37  |



|         |  |     |
|---------|--|-----|
| 3.3.4   | Kravanalyse .....                                      | 38  |
| 3.3.5   | Kravdokumentasjon.....                                 | 39  |
| 3.3.6   | Krav-verifisering og validering .....                  | 39  |
| 3.4     | Steg 3 Design og utvikling .....                       | 39  |
| 3.5     | Steg 4 Demonstrasjon .....                             | 41  |
| 3.6     | Steg 5 Evaluering .....                                | 42  |
| 3.7     | Steg 6 Demonstrasjon og evaluering .....               | 43  |
| 4       | Krav for rammeverk .....                               | 45  |
| 4.1     | Forretningsforståelse .....                            | 47  |
| 4.2     | Dataforståelse .....                                   | 49  |
| 4.3     | Dataforberedelse.....                                  | 51  |
| 4.4     | Modellering.....                                       | 53  |
| 4.5     | Evaluering .....                                       | 55  |
| 4.6     | Implementering .....                                   | 57  |
| 5       | Skisse av rammeverk .....                              | 59  |
| 5.1     | Forretningsforståelse .....                            | 61  |
| 5.1.1   | Benefit Dependency Network – analyse.....              | 61  |
| 5.1.1.1 | Problemsentrert tilnærming.....                        | 61  |
| 5.1.1.2 | Innovasjonsbasert tilnærming .....                     | 62  |
| 5.1.2   | PESTEL-analyse .....                                   | 62  |
| 5.1.3   | Hvordan gjennomføre en PESTEL-analyse .....            | 63  |
| 5.1.4   | SWOT-analyse .....                                     | 64  |
| 5.2     | Dataforståelse .....                                   | 66  |
| 5.3     | Dataforberedelse.....                                  | 72  |
| 5.4     | Modellering.....                                       | 76  |
| 5.5     | Evaluering .....                                       | 78  |
| 5.6     | Implementering .....                                   | 81  |
| 6       | Diskusjon.....   | 85  |
| 6.1     | Forskningstilnærming .....                             | 85  |
| 6.2     | Valg og bruk av kilder.....                            | 85  |
| 6.2.1   | Primærkilder.....                                      | 86  |
| 6.2.2   | Sekundærkilder .....                                   | 87  |
| 6.3     | Drøfting av resultat .....                             | 87  |
| 6.3.1   | Drøfting av utvikling av kravene til rammeverket ..... | 90  |
| 6.3.2   | Drøfting av funnene i rammeverket .....                | 90  |
| 7       | Konklusjon .....                                       | 99  |
| 8       | Referanser .....                                       | 101 |

|     |   |     |
|-----|---|-----|
| 9   | Vedlegg .....   | 113 |
| 9.1 | Vedlegg 1: Resultat interessentanalyse .....                      | 113 |
| 9.2 | Vedlegg 2: Samtykkeerklæring .....                                | 116 |
| 9.3 | Vedlegg 3: Intervjuguide #1 .....                                 | 119 |
| 9.4 | Vedlegg 4: Intervjuguide #2 og #3 .....                           | 120 |
| 9.5 | Vedlegg 5: Intervjuguide #4 .....                                 | 121 |
| 9.6 | Vedlegg 6: Meldeskjema for behandling av personopplysninger ..... | 122 |
| 9.7 | Vedlegg 7: Vurdering av behandling av personopplysninger .....    | 125 |

## Figurer

|         |   |    |
|---------|---|----|
| Figur 1 | Design Science Research Methodology. Modellen er laget basert på Peffers modell for DSRM (Peffers et al., 2014) ..... | 33 |
| Figur 3 | Vår tilnærming til DSRM .....   | 34 |
| Figur 4 | Rammeverk for tekniske og organisatoriske krav for utvikling av data-drevne makroanalyser .....                       | 60 |

## Tabeller

|          |   |     |
|----------|---|-----|
| Tabell 1 | Intervjuobjekt .....                            | 38  |
| Tabell 2 | Første forslag til utforming av rammeverk ..... | 41  |
| Tabell 3 | Formål ved en PESTEL-analyse .....              | 62  |
| Tabell 4 | Roller i dataforståelse .....                   | 70  |
| Tabell 5 | Resultat av interessentanalyse .....            | 114 |

# Forkortelser

|          |   |
|----------|---|
| BI       | Business Intelligence   |
| BDN      | Benefit Dependency Network  |
| CRISP-DM | CRoss Industry Standard Process for Data Mining                     |
| CRM      | Customer Relationship Management                                    |
| DS       | Design Science  |
| DSRM     | Design Science Research Methodology                                 |
| ERP      | Enterprise Resource Planning  |
| ETL      | Extract, Transform and Load   |
| HDFS     | Hadoop Distributed File System                                      |
| IEEE     | Institute of Electrical and Electronics Engineers                   |
| IoT      | Internet of Things  |
| IT       | Informasjons Teknologi  |
| KDD      | Knowledge Discovery in Databases                                    |
| MDRE     | Marked-Driven Requirements Engineering                              |
| PESTEL   | Political, Economic, Social, Technological, Environmental and Legal |
| SEMMA    | Samle, Explore, Modify, Model and Assess                            |
| SWOT     | Strengths, Weaknesses, Opponents, Threats                           |

# 1 Innledning

## 1.1 Bakgrunn

Utgangspunktet for denne oppgaven er å utforske hvordan virksomheter kan bruke data-drevne metoder til å håndtere den utfordrende situasjonen vi ser i verdensbildet i dag. Mange virksomheter opplever at verdenssituasjonen skaper stor uforutsigbarhet. De har bare de siste årene opplevd utfordringer knyttet til covid-pandemien, krig i Europa, drastiske prisøkninger, klimautfordringer, det grønne skiftet, med mer. Oppgavestiller twoday Avento ønsket at vi skulle utforske hvordan virksomheter kan ta i bruk data-drevne metoder for å møte disse utfordringene.

I kjølvannet av covid-pandemien ble digitalisering et gjennomgripende utviklingstrekk som berørte alle virksomheter i Norge i en eller annen grad. På grunn av varierende ressurser og innværende kunnskap i virksomheter, har de forholdt seg forskjellig til hvordan de skulle gjennomføre digitalisering. Mange av disse bedriftene er allerede i gang med digitalisering på forskjellig nivå og med forskjellig kompleksitet. De aller fleste vil ta i bruk ny teknologi og digitale løsninger de kommende årene for å lykkes videre (Innovasjon Norge, 2021). Dette krever at virksomheter må bli i stand til å løse oppgaver og overkomme utfordringer i konkrete settinger, basert på digital kunnskap, ferdigheter og holdninger. Likevel er ressursbegrensninger og manglende struktur for strategisk ledelse, utfordringer virksomheter ofte støter på. Dette begrenser evnen de har til å utnytte markedsmuligheter og ny teknologi.

Den pågående konflikten mellom Russland og Ukraina, viser hvordan både samfunn som helhet, men også virksomheter, blir påvirket av omstendigheter i verdenssituasjonen. En studie av Lim et al. fra 2022, så på både samfunn og virksomheter i og utenfor de to krigsinvolverte landene. Av faktorer som påvirker samfunn innad i krigsinvolverte land, konkluderte studien med tre hovedfaktorer. Den første er *begrenset adgang til grunnleggende nødvendigheter og pengeressurser*, den andre er *økt arbeidsledighet og redusert kjøpekraft*, og den tredje er en *økning i flyktninger og asylsøkere fra de krigsinvolverte landene*. Av faktorer som påvirker samfunn utad krigsinvolverte land, konkluderer forfatterne med de to faktorene *forsyningsmangel og inflasjon*, samt *trusselen desinformasjon utgjør*. Når det kommer til virksomheter innad i krigsinvolverte land, konkluderes det med tre hovedfaktorer her også. Den første er *trusselen av cyberangrep*, den andre er *trussel mot digital og bærekraftig vekst*, og den tredje er *kortvarige og langvarige sanksjoner og bistandsinnsatser*. Til slutt presenterer studien to hovedfaktorer som påvirker virksomheter utad de krigsinvolverte landene. Disse to er *utfordringer knyttet til etiske og moralske forpliktelser*, samt *utfordringer knyttet til merkevareledelse* (Lim et al., 2022). Overnevnte er bare et eksempel på hvordan geografisk begrensede omstendigheter kan ha store konsekvenser på virksomheter også utenfor direkte involverte land, som et resultat av en globalisert verden understøttet av felles interesser og avhengigheter. Med tanke på hvordan virksomheter kan møte utfordringer i verdenssituasjonen, står analyser av makroomgivelsene sentralt.

Desto mer komplekst verdensbildet blir, jo vanskeligere blir det for mange virksomheter å ta riktige beslutninger. Dette gjelder spesielt dersom informasjonsgrunnlaget man baserer beslutninger på er svakt. Virksomhetsoperasjoner akselererer, og gode

beslutninger må i større grad tas, i hurtigere tempo. Økende data-tilfang og avanserte analyseverktøy bidrar til at virksomheter har mulighet til å ta mer data-drevne beslutninger. Virksomheter velger å benytte analytiske applikasjoner til å forbedre effektiviteten av forretningsprosesser, øke synligheten, lønnsomheten, og øke produktiviteten og resultatene. Virksomheter blir drevet mot analytiske løsninger på grunn av behovet for å forbedre beslutningsprosesser. De møter også tidsbegrensninger når det gjelder å levere etter forretningskrav for rask beslutningstaking i konkurranseintensive og raskt endrende markeder (Nucleus Research, 2014).

## 1.2 Problemstilling og forretningsproblem

Med grunnlag i bakgrunnen for oppgaven, utarbeidet vi problemstillingen:

*«Vi vil finne ut hva som kreves teknisk og organisatorisk av et rammeverk for utvikling av data-drevne makroanalyser. Med bakgrunn i dette ønsker vi å forsøke å skissere et rammeverk.»*

Forretningsproblemet er forankret i at et flertall av virksomheter ikke klarer å realisere mange av de verdiene som ligger i bruk av data for å drive analyse av makroomgivelser. En studie av Berndtsson et al. (2020), undersøkte 13 forskjellige virksomheters forsøk på å bli data-drevne gjennom hvordan de strategisk gikk frem for å gjennomføre prosjektene, samt hvilke hindre de støtte på. Ifølge to studier beskrevet i rapporten, har bare 30% av virksomhetene som forsøkte, faktisk lyktes med å bli data-drevne (Bean & Davenport, 2019; Halper & Stodder, 2017. sitert i Berndtsson et al., 2020). Studien til Berndtsson et al. viser at de fleste virksomheter starter sin reise med et pilotprosjekt, velger å ta snarveier når de utvikler strategier, og støter på tidligere rapporterte hindringer. I tillegg avdekket de at virksomheter ikke bruker en helhetlig endringsledelsesprosess. Rapporten beskriver at en mulig løsning på dette, er å utvikle strategier som tar for seg både tekniske og ikke-tekniske problemer samtidig. Dette begrunnes med at det ikke ser ut til å bli gjort av mange virksomheter i dag. Flere av hindrene virksomheter støter på, forklares at lett kunne vært unngått eller redusert dersom veletablerte strategier hadde vært til stede. Videre beskriver de at virksomheter som ønsker å videreutvikle bruken av data-drevne analyser, burde bruke metoder fra endringsledelse som veiledende for prosjektene. Dette i stedet for å styre prosessen ad-hoc.

Det virker som at de fleste virksomheter forstår *hvorfor* de burde bli data-drevne. Likevel tyder den relativt lave graden av transformasjonssuksess på at spørsmål knyttet til *hva* det innebærer og *hvordan* man skal bli det, i stor grad forblir ubesvart. Ifølge en rapport utarbeidet av MIT Sloan Management Review og IBM Institute for Business Value, er det største hinderet for å oppnå konkurransefortrinn ved bruk av data/stordata, mangel på forståelse for hvordan man bruker analyse av dataen til å forbedre virksomheten (Lycett, 2013). Analyser av makroomgivelser vil være kritiske for å danne en forståelse av verdenssituasjonen virksomheter kan bruke som beslutningsgrunnlag i operasjonelle, taktiske og strategiske beslutningsprosesser. Det ligger potensielt mye verdi i analyser basert på tilgjengelig, men frem til nå ubrukt data i mange virksomheter. Vi ønsker å svare på behovet Berndtsson et al. (2020) konkluderte med i rapporten deres, hvilket også reflekteres i flere av kildene nevnt tidligere. Dette ønsker vi å gjøre gjennom å utvikle et rammeverk som tar for seg både de tekniske og organisatoriske faktorene, samt sammenhengene mellom disse. Dette må adresseres av virksomheter som ønsker å benytte data-drevne analyser av makroomgivelser. Spesifikt med tanke på bruk av data-drevne analyser av makroomgivelser, eksisterer det lite litteratur som omhandler

fagfeltet basert på våre litteratursøk. Vi ønsker å besvare *hva* det innebærer for virksomheter, dvs. hvordan en virksomhet som benytter data-drevne analyser av makroomgivelser strukturelt kan se ut. I endringsledelseslitteratur er en ofte nevnt nøkkelfaktor, en etablert felles forståelse av den nåværende situasjonen, samt en forståelse av ønsket slutttilstand. For å danne sistnevnte, er det kritisk å forstå hva et potensielt digitaliseringsprosjekt innebærer. Først da kan man formulere en strategi for hvordan man skal oppnå ønsket slutttilstand. Gjennom rammeverket ønsker vi å besvare hva som kreves teknisk og organisatorisk internt i en virksomhet, for at de videre skal kunne utvikle bruk av data-drevne makroanalyser.

### 1.3 Avgrensninger

Besvarelsen tar for seg hvordan virksomheter kan bruke data-drevne metoder til å håndtere den utfordrende situasjonen vi ser i verdensbildet i dag. Oppgavens resultat vil være basert på informasjon samlet inn og prosessert fra oppgavestiller, samt eksisterende litteratur og informasjon. Likevel vil oppgaven være relevant for andre virksomheter ettersom vi ser på løsninger for å ta i bruk data-drevne makroanalyser. Rammeverket som skisseres i denne oppgaven er derfor relevant og informerende overfor virksomheter som søker informasjon rundt data-drevne metoder, makroanalyser, eller som ønsker en løsning for effektivisering av beslutningstaking knyttet til makroomgivelser.

### 1.4 Oppgavens oppbygging

|                                  |  |
|----------------------------------|--|
| Kapittel 2 - Teoretisk grunnlag  | Kapittelet vil gi en innføring i teorien som ligger til grunn for oppgaven.  |
| Kapittel 3 - Metode              | Tar for seg metodene brukt i oppgaven. En beskrivelse av informasjonsinnhenting som er benyttet. I tillegg følger en gjennomgang av aktivitetene i metoden |
| Kapittel 4 – Krav til rammeverk  | Her fremlegges funn og resultater etter de fire første stegene i metoden.  |
| Kapittel 5 – Skisse av rammeverk | Her presenteres skisse av rammeverket.   |
| Kapittel 6 - Diskusjon           | Inneholder en drøfting av metode og resultatene oppnådd i kapittel 4 og 5, samt <u>ulike</u> styrker og svakheter ved rammeverket.                         |
| Kapittel 7 - Konklusjon          | Inneholder en helhetlig oppsummering av prosjektet, og en besvarelse av problemstillingen.   |

## 1.5 Om oppgavestiller

Denne besvarelsen utarbeides i samarbeid med twoday Avento. Dette er et konsultentselskap med spisskompetanse innenfor IT-rådgivning, systemutvikling, Business Intelligence (BI) og digital markedsføring. Med over 90 konsulenter fordelt på kontorene i Ålesund, Ørsta og Oslo, har de lang erfaring med design og utvikling av forretningskritiske datasystemer. Kundeporteføljen er spredt over hele Norge og deler av Sverige og kundene er alt fra mindre, lokale næringsdrivende til store internasjonale konsern (*twoday Avento*, u.å.).

Twoday Avento er en del av konsernet twoday som har 2200 ansatte i Norge, Sverige, Danmark, Finland, og Litauen, med hovedkontor i København. Twoday ble etablert i 2022 da 22 tidligere Visma selskaper fra divisjonen Visma Custom Solutions ble skilt ut fra Visma. Med sine 8000 kunder, omsatte selskapet for om lag 280 millioner euro i 2022 (twoday, u.å.). Konsernets hovedfokus er å bidra med digital transformasjon blant deres kunder. Dette innebærer å utvikle nyttige løsninger som gjør kundenes hverdag enklere. Blant disse løsningene finner man eksempler som effektivisering, automatisering, og optimalisering gjennom en digitaliseringsprosess av de allerede eksisterende produktene og tjenestene.

## 2 Teoretisk grunnlag

### 2.1 Datavitenskap

*Datavitenskap (eng. Data Science) er læren om generaliserbar utvinning av kunnskap fra data* (Dhar, 2013). Dette er et tverrfaglig felt som bruker ustrukturerte eller strukturerte data for å gi innsikt i, eller kunnskap om et område, ved hjelp av vitenskapelige metoder, analyse, algoritmer og fremgangsmåter. I motsetning til tradisjonelle virksomhetsanalyser, er datavitenskap et fenomen som står nært maskinlære, modellering, statistikk, algoritmer, analyser og matematikk. For å skape størst mulig verdi av virksomhetens tilgjengelige data, kombinerer datavitenskap de nevnte fenomenene (Data Science Association, 2020). Dermed vil en virksomhets kunnskap, samt handlingskraft øke, som potensielt kan gi de et bedre konkurransefortrinn i markedet de opererer innen.

Et underpunkt i datavitenskapen er data-drevne metoder. Dermed vil datavitenskap være i stand til å synliggjøre og predikere potensielle trender og mønstre som foreligger i data, som skaper stor verdi for brukeren (O'Neil & Schutt, u.å.). Virksomheter kan oppleve verdier som bedre beslutningsstøtte, og bedre utviklinger av tjenester og produkter. Sistnevnte begrunnes med at all den tilgjengelige dataen kan avsløre mulige mangler i et marked, eller fremtidige trender innenfor sitt marked. Samtidig vil datavitenskap også bidra til automatisering, optimalisering og effektivisering i områder der det foreligger et behov. Analyser av den tilgjengelige dataen vil bidra til å evaluere allerede eksisterende løsninger, samt utvikle nye løsninger som virksomheten vil tjene på (O'Neil & Schutt, u.å.).

For å anvende datavitenskap på en suksessfull måte, vil man være avhengig av enkelte forutsetninger. Bred og generell kompetanse rundt fagfeltet vil være en fundamental forutsetning for å lykkes. Dette innebærer særlig kompetente dataforskere med en god disiplin, som gjerne har en bakgrunn i statistikk, databehandling, og en god formidlingsevne, slik at man kan formidle resultatene på en strukturert måte (Martin, 2020). Samtidig vil tilgjengelig data være en nøkkelfaktor, da dataen legger grunnlaget for å kunne drive analyser. Det er viktig å ta i betraktning at datavitenskap befinner seg i en dynamisk utvikling, og vil dermed kontinuerlig endres parallelt med tiden (Martin, 2020). Derfor bør man være i stand til å foreta raske endringer og justeringer, dersom situasjonen skulle kreve det.

### 2.2 Data

For å utføre datavitenskap, må man inneha data. All data starter som rådata, ofte også omtalt som kilde- eller primærdata. Dette er data som ikke er behandlet, kodet, formatert eller analysert for nyttig informasjon. Rådata kan være en verdifull ressurs, men er vanligvis vanskelig å forstå eller handle på, da den er visuelt rotete og mangler sammenheng. Rådata er samlet inn fra én eller flere kilder, men er fortsatt i sin opprinnelige, uendrede tilstand. På dette stadiet, avhengig av innsamlingsmetoden, kan dataene inneholde mange menneskelige, maskinelle eller instrumentelle feil, eller mangle validering (Ot, 2023).



Når rådata er behandlet, omtales det som data. Før rådata kan benyttes til å gjennomføre analyser, må den som regel behandles på en eller flere ulike måter, slik at det blir mulig å trekke ut nyttig informasjon fra den (Ot, 2023). Dette kan for eksempel være å strukturere, organisere, eller renske dataen.

Når man henviser til brukbarheten av et gitt datasett (en samling av data) og dets egenskaper for et valgt brukstilfelle, omtales dette ofte som *datakvalitet*. Datakvalitet vil dermed avhenge av hvem som benytter dataen, og hva det benyttes til. Et valgt sett med data kan være svært passende for et brukstilfelle, men lite passende grunnet dårlig datakvalitet for et annet (Nikiforova, 2018b, som sitert i Nikiforova, 2020). Dersom man forholder seg til flere brukstilfeller av samme data, vil det dermed ofte være nødvendig å definere ulike krav for hvordan datakvaliteten vurderes. Det vil heller ikke være mulig å oppnå et nivå av datakvalitet som tilfredsstillende alle mulige brukstilfeller (Nikiforova, 2020). Data Management Association UK foreslår og beskriver følgende seks tradisjonelle dimensjoner å vurdere datakvalitet etter (gov.uk, 2021):

**Nøyaktighet** (eng. accuracy) - Data er nøyaktig når det reflekterer virkeligheten. Dette kan for eksempel referere til korrekte navn, adresser eller representere faktiske og oppdaterte data. Informasjon fra den virkelige verden kan endre seg over tid. Dette gjør det utfordrende å opprettholde nøyaktigheten. Man bør jevnlig gjennomgå data som sannsynligvis vil endre seg over tid. Høy nøyaktighet i dataene gjør det mulig å produsere analyser man kan stole på, og det fører også til korrekt rapportering og mer sikre beslutninger.

**Fullstendighet** (eng. completeness) - Data anses som fullstendig når all data som er nødvendig for en bestemt bruk, er til stede og tilgjengelig for å brukes. Det handler ikke nødvendigvis om å sikre at 100% av datafeltene er fullstendige, men om å avgjøre hva som er kritisk data og hva som er valgfritt. Fullstendighet er ikke det samme som nøyaktighet, siden et fullstendig datasett fortsatt kan inneha feilaktige verdier.

**Unikhet** (eng. uniqueness) - Unikhet måler antall duplikater. Data er unik hvis den bare vises en gang i en datamengde. Det er en høy risiko for duplisering, dersom man kombinerer datamengder. Man burde dermed alltid sjekke dataene for unikhet.

**Samsvar** (eng. consistency) - Samsvar oppnås når dataverdiene ikke er i konflikt med andre verdier i et datasett eller på tvers av forskjellige datasett. Dersom for eksempel en person er registrert i to forskjellige datasett, med to ulike bostedsadresser, vil disse verdiene være i konflikt. Samsvarende data gjør det enklere å kombinere data fra flere kilder. Dette supplerer igjen dataen din og øker nytteverdien av den.

**Aktualitet** (eng. timeliness) – Aktualitet indikerer om dataene er tilgjengelige når de forventes og behøves. Avhengig av bruksområde, kan aktualitet bety ha varierende betydning. I forhold til den operasjonelle driften av et sykehus, vil det for eksempel være kritisk å sikre oppdatert og tilgjengelig data om forskjellige faktorer til enhver tid. Imidlertid kan det være akseptabelt å bruke tidligere tall fra helse rapporter for å forutsi omsorgsbehov og planlegge helse- og omsorgstjenester. Aktualitet kan være viktig fordi det gir verdi til informasjon som er spesielt tidsfølsom.

**Gyldighet** (eng. validity) - Gyldighet defineres som i hvilken grad dataene samsvarer med forventet format, type og rekkevidde. For eksempel må en e-postadresse inkludere en alfakrøll, og måned skal være mellom en og tolv. Det å ha gyldige data betyr at det er enklere å kombinere den med andre datakilder, for eksempel i automatiserte dataprogrammer. Det at dataen er gyldig, betyr derimot ikke at den nødvendigvis er

korrekt. Selv om en e-postadresse inkluderer alfakrøll, trenger ikke det å bety at e-postadressen nødvendigvis er korrekt.

Overnevnte er eksempler på anerkjente dimensjoner av datakvalitet. Mange tror at kvaliteten på data kun avhenger av nøyaktighet, og vurderer dermed ikke andre betydningsfulle dimensjoner som kan være verdifulle for å oppnå høyere datakvalitet (Sidi et al., 2012). Gjennom søk i eksisterende litteratur identifiserte Sidi et al. (2012) 40 mulige dimensjoner av datakvalitet. Alle disse vil ikke være relevante for enhver virksomhet, men vil variere basert på ulike brukstilfeller samt hva de ønsker å oppnå gjennom bruk av dataen.

## 2.3 Stordata

Stordata (eng. big data) er en betegnelse på svært store datamengder som har en stor, mer variert og kompleks struktur, med vanskeligheter knyttet til å lagre, analysere og visualisere dataen. Prosessen med å forske på store datamengder for å avdekke skjulte mønstre og sammenhenger, kalles for stordataanalyse (Sagiroglu & Sinanc, 2013). I henhold til 3V-modellen (Wieczorkowski & Polak, 2014), innehar stordata følgende tre karakteristikk:

**Volum** (eng. *Volume*) – Svært store mengder prosessert data. Dette er først og fremst knyttet til den stadig økende veksten i mengden digital data tilgjengelig i en form som potensielt kan brukes.

**Hastighet** (eng. *Velocity*) - Variabilitet og dynamikk i prosessert data. Dette gjelder spesielt problemene med data som endres og genereres raskt, spesielt data som skal brukes kort tid etter endring eller generering. Denne karakteristikken gjelder typisk for sensorisk og strømmet data, samt data relatert til analyse av atferd på internett, inkludert sosiale medier.

**Variasjon** (eng. *Variety*) - Variasjon i prosessert data. Denne karakteristikken kan være knyttet til et svært stort antall ulike typer strukturerte databaseattributter i en relasjonsdatabase. Fremfor alt gjelder det imidlertid en variasjon av ustrukturerte data, som bilder, videoopptak, lyd, håndskrevne notater og annen data som ikke er egnet for lagring i en klassisk relasjonsdatabase.

## 2.4 Dataanalyse

Dataanalyse omfatter prosessen med å omgjøre data om til brukbar informasjon. Typiske dataanalyseprosjekter kan deles inn i flere ulike faser. Data blir vurdert og valgt ut, rengjort og filtrert, visualisert og analysert, og resultatene av analysen blir til slutt tolket og evaluert (Coursera, 2023). I henhold til analyser gjort av virksomheter, finnes det hovedsakelig fire forskjellige typer analyse. Disse er deskriptive, prediktive, preskriptive, og diagnostiske analyser (Adobe, 2022).

**Deskriptiv analyse** undersøker hva som har skjedd tidligere. Denne brukes når man undersøker tidligere datasett for mønstre og trender. Dette er kjernen i de fleste virksomheters analyser, fordi det svarer på viktige spørsmål som for eksempel hvor mye virksomheten solgte for, og om den nådde spesifikke mål. Deskriptiv analyse fungerer ved å identifisere hva man ønsker å måle, for deretter å samle inn og analysere dataen. Deskriptiv analyse gjør om strømmen av data som virksomheten har samlet inn, til informasjon virksomheten kan handle etter, planlegge rundt og måle (Adobe, 2022).

**Prediktiv analyse** har som mål å forutsi sannsynlige fremtidige utfall og utarbeide prognoser ved hjelp av historiske data. Der deskriptiv analyse søker å besvare spørsmålet «*Hva har skjedd?*», vil man bruke prediktive analyser til å besvare «*Hva kommer til å skje?*». Prediktiv analyse forsøker å forstå hvordan trender vil utvikle seg i fremtiden for så å forsøke å forutsi mulige utfall. Prediktive analyser er en mer kompleks versjon av dataanalyse, basert på at den bruker sannsynligheter for å gjøre prediksjoner i stedet for bare å tolke eksisterende fakta (Adobe, 2022).

**Preskriptiv analyse** bruker data fra ulike kilder for å identifisere mulige fremtidige utfall og forutsi den beste løsningen. Preskriptiv analyse er den mest avanserte av de hittil diskuterte typene, ettersom den gir forslag til handlinger som kan foretas videre, i stedet for kun data. Ved hjelp av preskriptive analyser kan man si noe om hva som bør skje, ikke bare hva som kan skje. Bruk av preskriptiv analyse gjør det mulig å ikke bare komme med forslag til utfall, men også begrunne hvorfor de vil skje. Preskriptiv analyse forsøker også å forutsi virkningen av fremtidige beslutninger, inkludert virkninger disse beslutningene kan ha på ulike deler av virksomheten. Preskriptiv analyse er en kompleks prosess som involverer mange variabler og verktøy som algoritmer, maskinlæring og stordata (Adobe, 2022).

**Diagnostisk analyse** forsøker å forklare hvorfor utfallet ble som det ble. Dette er en mer avansert versjon av deskriptiv analyse, som ikke kun går inn på hva som skjedde, men også hvorfor. Diagnostisk analyse identifiserer trender eller mønstre fra fortiden og går deretter et skritt videre ved å forsøke å forklare hvorfor trendene oppstod slik de gjorde. Diagnostisk analyse kan være et nyttig verktøy for virksomheter som ønsker å gjenskape gode resultater og samtidig unngå negative. Gjennom bruk av diagnostisk analyse kan man etablere et bedre grunnlag for å ta beslutninger og utvikle strategier. Dersom man klarer å forklare hvorfor ting skjedde, vil man kunne unngå mye prøving og feiling i flere ulike virksomhetsprosesser (Adobe, 2022).

## 2.5 Dataanalyseteam

Dataanalyser kan være svært komplekse. Den fullstendige prosessen fra innsamling av data til ferdig resultat, krever ulik kunnskap om forskjellige områder. I virksomheter blir dermed dataanalyseteam ofte etablert for dekke all nødvendig kunnskap og ferdigheter som kreves for å utføre en dataanalyse. Dataanalyseteamet er de som arbeider med å utvikle dataanalyser fra start til slutt (Pratt, 2021). Det finnes noen felles faktorer som alle dataanalyseteam burde bestå av for å kunne fungere. Dette inkluderer en forståelse av matematiske, teknologiske og forretningsmessige aspekter ved suksessfull dataanalyse. Det er svært sjeldent å finne én person med tilstrekkelig kompetanse på alle områder. De fleste virksomheter vil dermed sette sammen team bestående av ansatte som sammen utfyller den nødvendige kompetansen. Mindre virksomheter med få ansatte, vil derimot noen ganger ty til generalister som tar av seg alle aktiviteter i en analyseprosess. Typiske dataanalyseteam består ifølge Narasimman (2023) og Pratt (2021) av følgende roller:

**Dataanalytikere** (eng. data scientist) er de som finner, analyserer og kombinerer datakilder, visualiserer data, og benytter metoder for å utvikle modeller for analytisk innsikt. De er kjent med hele prosessen rundt dataoppdagelse, og kan presentere og kommunisere innsikt og andre funn til de andre ansatte i dataanalyseteamet.

**Dataingeniører** (eng. data engineer) sørger for at data er tilgjengelig og klar for dataanalyseprosjekter. De er ansvarlige for å bygge systemer som gjør dette mulig, og

er også ansvarlig for å opprettholde dem. De har typisk bakgrunn innen programvareutvikling eller datavitenskap, som understøtter deres fokus på teknologisk infrastruktur, data-innsamling, forvaltning og lagring. De arbeider tett med dataanalytikere for å sikre datakvalitet, forberede dataen for analyse, og implementere modellene som utvikles i dataanalyseprosjekter.

**Dataarkitekter** (eng. data architect) har hovedansvaret for å designe og overse implementering av underliggende systemer og dataarkitektur som teamet benytter seg av. Noen ganger kan dataingeniører også utfylle denne rollen.

**Maskinlæringsingeniører** (eng. machine learning engineer) er de som utvikler, implementerer, overser og ivaretar algoritmer og modeller basert på maskinlæring og kunstig intelligens, dersom virksomheten benytter seg av dette.

**Forretningsanalytiker** (eng. business analyst) arbeider med å støtte arbeidet til dataanalytikere, gjennom å evaluere forretningsprosesser og oversette forretningskrav til krav for dataanalysen.

**Dataoversettere** (eng. data translator) arbeider som et knutepunkt mellom dataanalyseteam og de ansatte ansvarlige for den operasjonelle driften av virksomheten. De kan hjelpe til med å planlegge prosjekter, og oversette innsikten fra dataanalyse over til anbefalte forretningsbeslutninger.

Dataanalyseteam kan også struktureres på forskjellige måter (Narasimman, 2023; Pratt, 2021). De tre hovedmåtene å strukturere dataanalyseteam, er gjennom en sentralisert, desentralisert eller hybrid struktur.

I en *desentralisert* struktur arbeider teammedlemmene under forskjellige forretningsenheter. Det gjør det enklere å utvikle analyser mer spesifisert mot anbefalte beslutninger for de individuelle forretningsenhetene, men kan hindre strategisk bruk av data på tvers av organisasjonen. En desentralisert struktur kan også være mer ressurskrevende (Narasimman, 2023; Pratt, 2021).

I en *sentralisert* struktur derimot, arbeider dataanalyseteamet på et overordnet forretningsnivå, under en ledelse som fordeler arbeidsoppgaver fra hele virksomheten. Dette kan føre til at analysene fører til bedre innsikt fra et overordnet forretningsperspektiv, og at analysene kan resultere i anbefalinger mer tilpasset hele virksomheten enn kun spesifikke forretningsenheter. En sentralisert struktur kan også hindre at ansatte spesialiseres innen spesifikke forretningsområder (Narasimman, 2023; Pratt, 2021).

I en *hybrid* struktur administreres dataanalyseteamet sentralt, men teammedlemmene fordeles over spesifikke forretningsenheter, og er ansvarlige for å hjelpe enhetene med å nå sine mål for å ta data-drevne beslutninger (Narasimman, 2023; Pratt, 2021).

Når det kommer til ledelse av dataanalyseteam, er den mest synlige lederen ofte en CDO (Chief Data Officer) (Pratt, 2021). I følge en analyse av store virksomheter gjort NewVantage Partner i 2020, rapportere 65% av virksomhetene at hadde en CDO (NewVantage Partner, 2020; som sitert i Pratt, 2021). En CDO er vanligvis medlem av toppledelsen. CDO-er administrerer virksomhetens data og datavitenskaps-funksjoner på tvers av virksomheten. Generelt rapporterer CDO til Chief Technology Officer (CTO) eller Chief Executive Officer (CEO) (Griffin, 2008). En CTO har ansvaret for å overse tekniske aspekter ved en virksomhet, og sørge for at alle maskinwaresystemer, applikasjoner, og nettverk kan hjelpe virksomheten oppnå målene deres. En CEO er det høyeste utøvende

nivået i en virksomhet, og har det overordnede ansvaret for virksomhetens forretningsbeslutninger (N. Ahmed, 2022).

## 2.6 CRISP-DM

CRISP-DM står for Cross-Industry Standard Process for Data Mining, og er en standardprosess for planlegging, utvikling og implementering av data-drevne prosjekter. Prosessmodellen gir en strukturert tilnærming til datavitenskapprosjekter (IBM, 2011). CRISP-DM er med god margin er den mest brukte modellen for å designe og strukturere datavitenskapprosjekter (IBM, 2011; J. Saltz, 2022). CRISP-DM er også en fleksibel modell, som kan tilpasses alle typer datavitenskapprosjekter (IBM, 2021a). I motsetning til andre, mer tekniske prosessmodeller, har CRISP-DM også et stort forretningsmessig fokus. Modellen skisserer en håndfast metode man skal forholde seg til, illustrert gjennom seks steg, noe som enkelt geleider virksomheten gjennom modellen på en sikker måte. Ved å starte modellen med forretningsforståelse, tvinger man brukeren til å arbeide seg gjennom et essensielt steg som hyppig blir hoppet over eller nedprioritert av virksomheter. Dermed vil man tidlig gjøre seg opp forretningsbehov tilknyttet virksomheten, og deretter forstå formålet og hensikten ved å anvende en slik prosessmodell (J. S. Saltz, 2021). Dette gjør at man tidlig vil få en forståelse av det reelle problemet, og at dette blir adressert internt i virksomheten. Deretter vil man opparbeide seg et syn på hvilke data som er tilgjengelige, og som kan disponeres for å løse det aktuelle problemet. For å velge aktuell data trenger det å foreligge en evaluering som validerer dataene som skal brukes. Her inngår også et valg av den mest passende analysemetoden. Avslutningsvis vil man også sikre seg en implementering, samt et stort fokus på vedlikehold og overvåkning gjennom siste steget av prosessmodellen.

Videre har modellen et sett unike egenskaper, som gjør at den skiller seg ut fra andre kjente prosessmodeller. I tillegg til å være en fleksibel prosessmodell, er CRISP-DM svært enkel å forstå, da stegene i stor grad er selvforklarende, samt bygger på logiske prinsipper. Ved at CRISP-DM er enkel å forstå, gjør dette at utenforstående enkelt kan tilegne seg en forståelse for hvordan prosessmodellen fungerer, og hva som må gjøres for å lykkes (J. S. Saltz, 2021). Dermed vil virksomheten som en enhet, klare å utforme en universell forståelse av modellen og dens formål. Dette gjør at virksomheten blir mer smidig, da man effektivt vil kunne arbeide på tvers av teams og avdelinger. Basert på det faktum at CRISP-DM i stor grad er tilpasningsdyktig, vil virksomheten også ha muligheten til å drives mer ressurseffektivt (J. S. Saltz, 2021). Ved å ha en forholdsvis enkel modell å forholde seg til, kan man med større sannsynlighet unngå å støte på store problemer og kontroverser, samtidig som man slipper å gjøre omfattende endringer innad i virksomheten. Det vil derfor være rimelig å anta at disse fordelene har bidratt til at CRISP-DM har såpass høy oppslutning blant prosessmodellene anvendt av virksomheter. Under vil vi kort forklare de seks hovedfase i livssyklusen til et typisk datavitenskapprosjekt som CRISP-DM definerer.

**Forretningsforståelse** er den første fasen. Her er målet å forstå hva virksomheten forventer å tjene på datavitenskap. Denne fasen involverer først å etablere/forstå forretningsmålene, og hvordan datavitenskap kan hjelpe virksomheten med å nå dem (IBM, 2021b). Deretter må man vurdere nå-situasjonen, noe som innebærer å blant annet forstå hvilke type data som er tilgjengelig, om virksomheten har nødvendig personell, hva som utgjør de største risikofaktorene for prosjektet, og om man har beredskapsplan klar for de ulike risikoene. Videre må spesifikke mål for

datavitenskapprosjektet etableres med grunnlag i forretningsmålene. Til slutt, må det legges en prosjektplan for datavitenskapprosjektet (IBM, 2021b).

**Dataforståelse** er den andre fasen. Denne fasen har som mål å undersøke den tilgjengelige dataen, hvilket som er kritisk for å unngå problemer i de følgende fasene. Her må man først samle inn nødvendig data fra de kildene som er tilgjengelig for virksomheten (IBM, 2021c). Deretter må dataens karakteristikk beskrives, med stor vekt på forklaring av dataens kvalitet og kvantitet. Videre burde datavitenskapsteamet utforske dataen, gjennom bruk av blant annet ulike visualiseringsverktøy. Dette kan gi en bedre forståelse av målene for prosjektet, hjelpe med formuleringen av hypoteser, og danne et bedre grunnlag for tredje fasen. Til slutt må dataens kvalitet verifiseres. Dette innebærer å undersøke potensielle feilkilder i dataen, som for eksempel manglende verdier, inkonsistens, feil i dataen, etc. (IBM, 2021c) .

**Dataforberedelse** er den tredje fasen i CRISP-DM. Denne fasen tar for seg alle aktiviteter som kreves for å konstruere det endelige datasettet som i neste fase skal benyttes i modelleringsfasen, ut ifra innledende rådata (Wirth & Hipp, 2000). Dette er en av de viktigste, men også mest tidskrevende fasene av datavitenskap (IBM, 2021d). Denne fasen tar vanligvis opp 50-70% av prosjektets gjennomføringstid. Godt forarbeid i de foregående fasene kan være med på å korte ned denne tiden. Det første steget i fasen er å velge ut hvilke data som er relevant. Generelt gjøres dette ved å enten velge ut rader eller kolonner i datasett. Deretter må man undersøke problemene i dataen nærmere, og rense den for kvalitetsproblemer. Videre må man ofte integrere data, dersom man har data fra flere kilder som skal brukes for å besvare de samme forretnings spørsmålene. Til slutt må dataen ofte formateres, avhengig av modelleringsteknikkene man vil benytte i neste fase(IBM, 2021d).

**Modellering** er den fjerde fasen. Her utarbeider man modeller basert på de endelige datasettene. Dette innebærer først å velge ut modelleringsteknikk typisk basert på type data, mål for prosjektet, og krav for å ta i bruk de ulike teknikkene (IBM, 2021e). Videre vil man ofte utvikle et test-design for å teste resultatene av modellen. Her vil man ofte dele dataen inn i trenings-, test- og valideringssett. Videre er neste steg å bygge selve modellen. Det er vanlig å bygge flere ulike modeller, før man i siste steg velger ut de mest passende basert på målet for datavitenskapprosjektet (IBM, 2021e).

**Evaluering** er den femte fasen. Her burde resultatene evalueres opp mot målene for datavitenskapprosjektet etablert innledningsvis i prosessen, slik at man sikrer at virksomheten faktisk kan ha nytte av resultatene (IBM, 2021f). Prosessen som helhet, burde også evalueres. Teamet burde reflektere over hvordan de ulike fasene i prosjektet ble gjennomført, og forsøke å avdekke muligheter for forbedring. Når alt dette er gjort, kan man vurdere om man skal gå videre til siste fase, eller gå tilbake i prosessen og gjøre endringer på modellene dersom resultatene ikke ble optimale (IBM, 2021f).

**Implementering** er den siste fasen i CRISP-DM. Implementering er prosessen med å benytte resultatene av prosjektet til å forbedre virksomheten. Dette kan for eksempel være å integrere modellene med virksomhetens databaser for å automatisere analyse av innkommende data, eller ved å benytte innsikten til å direkte fremme endringer i virksomheten (IBM, 2021g). Det vil være viktig å legge en plan for hvordan implementeringen skal gjennomføres. Dette vil innebære hvordan resultatene skal kommuniseres, på riktig måte, til de riktige personene. Dersom modeller skal integreres i andre systemer, vil dette kreve en plan for nøyaktig hvordan det skal gjennomføres. For disse modellene vil det også være viktig å planlegge hvordan de skal overvåkes og

vedlikeholdes, for å forsikre seg om at de fremdeles produserer gunstige resultater i henhold til blant annet validitet og nøyaktighet. Man burde også avgjøre hvilke faktorer som indikerer at modellen ikke presterer godt nok lenger, og hva som da eventuelt burde gjøres. Noen ganger holder det med små endringer i modellen, andre ganger må man starte et nytt datavitenskapprosjekt og bygge en ny modell fra bunn av. Helt til slutt bør man i denne fasen produsere en sluttrapport for prosjektet, samt gjennomføre en avsluttende prosjektgjennomgang (IBM, 2021g).

## 2.7 Kunstig intelligens

Kunstig intelligens (KI) (eng. Artificial Intelligence (AI)) er digitale systemer som utfører oppgaver som normalt krever menneskelig intelligens. Dette gjør det mulig å automatisere oppgaver og gi mennesker støtte til å ta gode beslutninger (SINTEF, u.å.). EUs ekspertgruppe definerer kunstig intelligens på følgende måte:

*«Kunstig intelligente systemer utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Enkelte KI-systemer kan også tilpasse seg gjennom å analysere og ta hensyn til hvordan tidligere handlinger har påvirket omgivelsene.»* - (Regjeringen, 2020, s. 9)

Systemer basert på kunstig intelligens kan tolke data fra fysiske objekter som kameraer, sensorer, mikrofoner, etc. Slike systemer kan også samle inn data fra andre informasjonskilder, som tilkoblede databaser. Systemet vil deretter analysere dataene, ta beslutninger, og utføre handlinger. I noen typer systemer er det en tilbakemeldingsløyfe, som gjør at systemet lærer basert på tilbakemelding fra bruker og/eller egne erfaringer (Regjeringen, 2020). Kunstig intelligens er et paraplybegrep som dekker alle former for intelligente systemer. Det er et tverrfaglig forskningsområde bestående av både datavitenskap og informatikk (Helo & Hao, 2021). Begrepet kunstig intelligens dekker en rekke ulike tilnærminger og teknikker, som maskinlæring, maskinresonnering, og enkelte metoder innen robotikk (Regjeringen, 2020).

Maskinlæring (eng. machine learning) er muligens den viktigste teknologien dekket av begrepet kunstig intelligens. Maskinlæring omhandler maskiners evne til å kontinuerlig forbedre ytelsen, uten at mennesker må gripe inn for å programmere nøyaktig hvordan de skal utføre alle oppgavene de blir gitt. Maskinlæring gjør det mulig for maskiner å behandle data, identifisere mønstre i store datasett, ta beslutninger basert på tilbakemelding, og deretter utføre oppgaver på egenhånd. Bare i løpet av de siste årene har maskinlæring blitt langt mer effektivt og tilgjengelig (Brynjolfsson & McAfee, 2017).

## 2.8 Data-drevne metoder

Data-drevne metoder er kjent som bruken av datavitenskapelige teknikker og anvendelsen av dataanalyser, for å bedre kunne ta informerte beslutninger (Patil & Mason, 2015). Dermed er det nødt til å foreligge et datagrunnlag som kan disponeres til analyser. Samtidig er brukeren av data-drevne metoder nødt til å inneha aktuelle teknologier, verktøy, ressurser og kompetanse rundt fagfeltet. Dette er essensielt for å lykkes med målene man setter seg ved bruken av metodene. I et forretningsperspektiv kan denne utnyttelsen av data potensielt gagne bedrifter, da analyser kan f.eks. ved hjelp av maskinlæring avsløre ulike mønstre i dataen, noe mennesker ikke alltid er i stand til å gjøre. Som et resultat av mønstrene, vil man med større sannsynlighet kunne være i stand til å forutsi fremtidige utfall og trender. Dette kan bidra til å opparbeide seg et konkurransefortrinn i markedet (Kar & Dwivedi, 2020). Et verktøy som kontinuerlig

forvalter den tilgjengelige dataen, kan derfor sees på som en rimelig og stabil løsning for virksomheter.

Annet enn å gi bedre beslutningsgrunnlag, kan data-drevne metoder bidra til effektivisering, optimalisering og automatisering av allerede eksisterende prosesser. Gjennom en automatisering av tradisjonelle prosesser, vil effektiviteten øke (Kar & Dwivedi, 2020). I tillegg til automatisering, kan man optimalisere prosesser. Optimalisering av en prosess vil gjøre en prosess mer brukervennlig, samt gi brukeren et bedre utbytte. I likhet med automatisering, vil optimalisering også bidra til en økt effektivitet, både tids- og ressursmessig. Tidligere nevnte vi at data-drevne metoder kan avsløre skjulte mønstre gjennom bl.a. maskinlæreteknikker, noe som resulterer i en prediksjon av fremtidige utfall og trender. Basert på et tilsvarende grunnlag, kan en slik analyse gi en oversikt over hvilke muligheter man har tilgjengelig, og dermed gi mer informasjon til å fatte en beslutning. Gjennom grunnlaget som analysene opparbeider, vil man være i stand til å ta fortløpende vurderinger på endringer som man ønsker å gjennomføre. Samtidig kan man bli oppmerksom på tilgjengelige muligheter som tidligere ikke har vært kjent. Dersom man blir klar over om disse mulighetene, vil man være mer kapabel til å ta beslutninger som potensielt leder til innovasjon (Patil & Mason, 2015). Innovasjonen kan være internt i virksomheter, som for eksempel en optimalisering av en bestemt prosess. I tillegg kan innovasjonen bidra til endring i markedet, for eksempel gjennom et nytt produkt eller tjeneste.

## 2.9 Virksomheters bruk av data

I takt med en økning i mengden data generert globalt, øker også flere virksomheters evne til å benytte seg av den. For å opprettholde konkurranseevnen, vil dette ofte kreve at virksomheter omstiller seg. Begrepet «digital modenhet» blir ofte brukt i litteratur for å beskrive hvor mottakelig virksomheter er for digital endring. Det finnes ingen universalt akseptert definisjon av hva det innebærer, men basert på eksisterende litteratur kan man si at handler om hvordan virksomheter systematisk forbereder og tilpasser seg digital endring. Dette innebærer organisatorisk endring i form av ansatte og prosesser, men også teknisk endring i form av ny teknologi og hvordan dette integreres på tvers av virksomheten (Aslanova & Kulichkina, 2020).

Buitelaar (2018) identifiserte basert på eksisterende litteratur, åtte felles dimensjoner virksomheter som benytter seg av data består av (Buitelaar, 2018. s. 33). Disse beskrives som følgende:

- **Data** – Data er drivstoffet for alle data-drevne aktiviteter. Hvordan skaffer og administrerer virksomheten data?
- **Målinger (eng. metrics)** - Nøkkelen til å måle produksjon og administrere ytelse. Hvordan bruker, samler og beriker virksomheter nøkkeltallsindikatorene sine?
- **Ferdigheter (eng. skills)** - Essensielt for å drive en datadrevet virksomhet. Ansetter og utdanner virksomheten de riktige menneskene med de riktige ferdighetene?
- **Teknologi** - Fundamentet for en datadrevet virksomhet. Hvilken teknologi trenger virksomheten for å bygge en analytisk prosess?
- **Ledelse (eng. leadership)** - Hjørnesteinen for en vellykket analytisk transformasjon. Hvordan styrer ledelsen transformasjonen på en vellykket måte?
- **Kultur** - Kraften bak en datadrevet virksomhet. Hvordan påvirker og fremmer kulturen en datadrevet tilpasning?



- **Strategi** - Planen for suksess. Hvilken rolle har analyse i planene og visjonene for virksomhetens fremtid?
- **Smidighet (eng. agility)** - Evnen til å tilpasse seg og levere. Hvor godt er rollene og prosessene innad virksomheten organisert for å endre seg og levere?

Digitale endringer vil ofte kreve gjennomgående omstillinger i virksomheter, for at de skal være i stand til å realisere verdi fra dem. Ulike endringer kan påvirke alle de overnevnte åtte dimensjonene, hvilket ofte kan gjøre dem komplekse å gjennomføre, og i verste fall fører til at endringene mislykkes.

Mange virksomheter ytrer et ønske om å bli data-drevne. Til tross for dette, sliter de med å gjøre de omstillingene som kreves for å kunne bli det, samt realisere verdien av det. Dette skyldes ulike rapporterte barrierer på tvers av virksomhetene (Halper & Stodder, 2017). En virksomhet kan klassifiseres som data-drevet «*når den bruker analyse og data for å hjelpe med å drive handlinger, selv om handlingen er bevisst passiv*» (Halper & Stodder, 2017, s. 5). I teorien, kan data-drevne virksomheter benytte data i alle typer analyser, dvs. deskriptive, prediktive, preskriptive og diagnostiske, samt alle typer beslutninger, dvs. operasjonelle, taktiske og strategiske. I praksis kan man anta at de fleste virksomheter ønsker en kombinasjon av analyser og data-drevne beslutninger. Nucleus Research fant i en av deres studier, at gjennomsnittlig avkastning på investering per USD brukt på dataanalyse i 2014 lå på 13.01 USD (Nucleus Research, 2014). Dette tyder på at det ligger ekstremt mye verdi i bruk av dataanalyser for virksomheter. Et av hovedfunnene i en rapport utarbeidet av MIT Sloan Management Review og IBM Institute for Business Value, var at de best presterende virksomhetene rapporterte effektiv bruk av analyser som en nøkkelfaktor for differensiering. I tillegg til å understreke en allerede utbredt tro på at dataanalyse tilbyr verdi for virksomheter, fant rapporten ut at de best presterende virksomhetene fattet beslutninger basert på nøyte dataanalyser i mer enn dobbelt så stor grad som de lavest presterende virksomheten (Lavelle et al, 2010, sitert i Lycett, 2013).

En rapport utarbeidet av SAS Institute og MIT Sloan Management Review, viser at selv om optimismen rundt potensialet dataanalyse utgjør er høy, så synker graden av virksomheter som rapporterer konkurransefordeler som resultat av dataanalyser (Ransbotham et al., 2016, s. 5). Rapporten viser også at prosessen med å oppnå konkurransefordeler som resultat av analyser, krever et vedvarende engasjement for å endre rollen data utgjør i virksomhetens beslutningsprosesser. Virksomheter som opplever suksess gjennom dataanalyse, har oftere en strategisk plan bak innsatsen deres for å ta det i bruk, og denne planen er oftere kompatibel med virksomhetens overordnede forretningsstrategi. Rapporten viser i tillegg at de fleste virksomheter ikke er forberedt på de store investeringene og gjennomgående endringene som er nødvendig for å oppnå vedvarende suksess gjennom bruk av dataanalyse (Ransbotham et al., 2016). Ledende virksomheter er i høyere grad i stand til å gjøre data om til verdifull innsikt, samt å benytte data i interne beslutningstakingsprosesser. Andre virksomheter faller bak og sliter med å realisere verdien dataanalyser kan gi (Henke et al., 2016). Rapporten viser i tillegg at de fleste virksomheter ikke er forberedt på de store investeringene og gjennomgående endringene som er nødvendig for å oppnå vedvarende suksess gjennom bruk av dataanalyse (Ransbotham et al., 2016). Ledende virksomheter er i høyere grad i stand til å gjøre data om til verdifull innsikt, samt å benytte data i interne beslutningstakingsprosesser. Andre virksomheter faller bak og sliter med å realisere verdien data-drevne analyser kan gi (Henke et al., 2016). Basert på studiene ovenfor er det et relativt stort gap mellom fordelene dataanalyser kan gi virksomheter,

og hvor viktig virksomheten ser på bruken av det er, til hvor gode mange virksomheter er på å bruke dataanalyser i praksis.

## 2.10 Benefit Dependency Networking

Når det kommer til IT-investeringer, fokuserer de fleste organisasjoner på å implementere teknologien i stedet for å realisere de forventede gevinstene. IT-gevinststyring kan defineres som «*prosessen, organisering og styring slik at de potensielle fordelene ved bruk av IT faktisk realiseres*» (Peppard & Ward, 2007, s. 3). Benefit Dependency Networking (BDN) viser hvordan hver av gevinstene kan oppnås ved en kombinasjon av forretningsendringer og nye IT-funksjoner. Denne prosessen begynner gjerne med at IT-fagfolk og bedriftsledere sammen svarer på syv spørsmål om en potensiell IT-investering. Disse spørsmålene vil omtales i resultatkapittelet. Målet er å avdekke tre hovedområder ved en IT-investering:

- Målene – målytelsesforbedringer
- Måtene – måtene virksomheten må jobbe annerledes på
- Midlene – de muliggjørende IT-evner.

BDN skal også klargjøre hvem som må stå ansvarlig for å gjennomføre de spesifikke endringene og levere gevinstene. Nettverket skal gi grunnlaget for å utvikle både en business case for investeringen og en endringsplan for å levere gevinstene. Målet vil være bedre investeringsbeslutninger og gevinst-drevne implementeringsplaner som sammen fører til å realisere flere fordeler fra IT-investeringer.

## 2.11 Strategisk analyse av omgivelser

For at virksomheter skal styres riktig, benyttes ofte ulike former for strategisk analyse. Strategisk analyse er det første grunnleggende steget i strategisk ledelse av virksomheter, og omfatter analyse av de faktorer som er relevant i forhold til omgivelsene en virksomhet opererer i (Ülgen & Mirze, 2007, som sitert i Yüksel, 2012, s. 52). En virksomhets interne omgivelser kan forklares som ressursene og kapabilitetene til virksomheten, mens virksomhetens eksterne omgivelser omfatter faktorer utenfor virksomhetens direkte kontroll. De eksterne omgivelsene er relevante og kan påvirke virksomheten, men vil være svært vanskelig eller umulig for virksomheten å kontrollere. En virksomhets eksterne omgivelser består av makroomgivelser og bransjeomgivelser. Bransjeomgivelsene omfatter omgivelsene i bransjen virksomheten opererer i (Yüksel, 2012). En bransje er en inndeling av en næring på lavere orden, og er en samling av virksomheter som arbeider med produkter eller tjenester av samme slag. (Store Norske Leksikon, 2023). Et vanlig verktøy for å analysere bransjeomgivelsene, er femkraftsmodellen til Michael Porter. Denne modellen tar for seg følgende fem krefter på bransjenivå; kundenes forhandlingsstyrke, leverandørenes forhandlingsstyrke, trusler fra fremtidige konkurrenter, trusler fra nære substitutter, og til slutt konkurransesituasjonen (Porter, 2008). Makroomgivelsene til en virksomhet derimot, omfatter de politiske, økonomiske, sosiale, miljømessige, og juridiske faktorene som kan ha en direkte eller indirekte påvirkning på virksomhetens drift (Peppard & Ward, 2016, side 58).

En annen strategisk analysemetode er SWOT. SWOT-analyse brukes til å evaluere en organisasjons interne og eksterne faktorer. SWOT står for styrker (eng: Strengths), svakheter (eng: Weaknesses), muligheter (eng: Opportunities), og trusler (eng: Threats). Denne analysen anvendes av virksomheter for å identifisere både fordeler og ulemper ved produkter, tjenester, løsninger, eller virksomheten som helhet. Det er også

verdt å nevne at det kan brukes for å påvirke en beslutning, og dermed øke muligheten for suksess (Vikøren & Pihl, 2023). Ved å benytte seg av denne typen strategisk planleggingsmetode, vil virksomheter være i stand til å identifisere muligheter og trusler, og deretter utvikle en strategi for å utnytte mulighetene og redusere truslene. SWOT-analyser er svært ressurs og kostnadseffektive, da man vil være i stand til å gjennomføre en analyse med få ressurser. Resultatene kan benyttes og vurderes i virksomhetens overordnede strategiutvikling. Det er viktig å merke seg at SWOT-analysen ikke gir en fullstendig strategisk plan, men at det er et verktøy for å evaluere organisasjonens nåværende situasjon og identifisere muligheter og trusler. SWOT-analyser kan også brukes som en del av en større strategisk planleggingsprosess, hvor organisasjonen kan vurdere ulike alternativer og utvikle en strategi som tar hensyn til de interne og eksterne faktorene. For å styrke SWOT-analysen ytterligere, kan virksomheter også anvende andre strategiske verktøy som for eksempel PESTEL-analyse, Porters fem-krefter-modell, og verdikjedeanalyse. Dette er verktøy som kan bidra til at virksomheter evaluerer ulike faktorer som kan påvirke strategiske beslutninger, samt utvikle en omfattende strategi som tar hensyn til alle disse faktorene.

## 2.12 Makroanalyser

Makroanalyser er strategiske analyser av en virksomhets makroomgivelser. PESTEL-analyse er et verktøy utviklet for å forstå denne makrokonteksten en virksomhet opererer i. Her inngår faktorer som virksomheter flest har liten innvirkning på, men i stor grad kan bli påvirket av (Peppard & Ward, 2016, side 57-60). En grundig forståelse av disse faktorene er et viktig hjelpemiddel i god strategisk beslutningstaking for virksomheter i nærmest alle typer bransjer. Det vil hjelpe virksomheter å bedre forstå makroomgivelsene de opererer i, og hvordan de selv kan bli påvirket av dem. En slik forståelse er med på å danne et bedre beslutningsgrunnlag i de mange avgjørelsene en virksomhet kontinuerlig står ovenfor.

En analyse av makroomgivelsene er avgjørende for virksomheter dersom de ønsker å utvikle et bærekraftig konkurransefortrinn, og identifisere muligheter eller potensielle trusler (Yüksel, 2012). PESTEL-analyse utgjør to basis-funksjoner for en virksomhet. For det første, identifiserer det miljøet virksomheten opererer i, og bidrar til å danne en bedre forståelse av det. For det andre, tilfører det data og informasjon som muliggjør prediksjoner av situasjoner og omstendigheter som kan oppstå i fremtiden, noe som kan være essensielt i strategisk beslutningstaking (Yüksel, 2012).

Som tidligere beskrevet er makroomgivelser eksterne faktorer som ikke er like lett å påvirke som mikroomgivelser. PESTEL er svært vanlig å bruke i utviklingen av markedsstrategi. Faktorene PESTEL synliggjør vil være hensiktsmessig å ta med seg videre i utviklingen av en SWOT-analyse, som vi redegjør for senere. PESTEL står for *Politiske* (Political)-, *Økonomiske* (Economical)-, *Samfunnsmessige* (Sociological)-, *Teknologiske* (Technological)-, *Bærekrafts* (Enviromental)- og *Juridiske* (Legal)- faktorer.

**Politiske** faktorer kan for eksempel, være begrensninger eller muligheter i dagens politikk. Skatter og avgifter, samt samfunnsøkonomi har mye å si for hvordan du kan drive virksomheten din (Magnussen, 2020). Det vil også dekke kommende valg og fremtidige skrifter i politisk styre.

**Økonomiske** faktorer kan påvirke muligheten for handel, lokalt og globalt, samt at de har implikasjoner for digital markedsføring. Disse faktorene kan være priskonjunkturer, og om dagens økonomi tilsier at målgruppen kan benytte seg av ditt produkt eller ikke.

Dette inkluderer også inflasjon, arbeidsledighet, valutakurser, minstelønn og renter (Reding, 2021). Denne faktoren er viktig for forretningsbeslutninger.

**Samfunnsmessige** faktorer kan være ting som påvirker samfunnet, sosiale normer og trender som er i markedet. Det kan være kulturelle faktorer og trender som påvirker folks livsstil (Reding, 2021). Typiske faktorer er befolkningsvekst, helsebevissthet, helse og sikkerhet. Det er viktig for en virksomhet å holde seg oppdatert på disse og forhåpentligvis prøve å predikere hvilke trender som kommer, kommer for å bli eller kommer til å forsvinne. Dette for å unngå å trå feil. Hvilke trender er spesielle for din virksomhets næring og marked. Hva syns samfunnet om det din bedrift gjør (Magnussen, 2020)?

**Teknologiske** faktorer synliggjør hva din virksomhet burde gjøre for å holde seg oppdatert på teknologi. Teknologi transformerer samfunnet raskt og tvinger virksomheter til å innovere. I tillegg påvirker teknologiske fremskritt andre områder av denne analysen, eks. økonomiske faktorer. For en virksomhet kan dette bety faktorer som automatisering, programvareutvikling og teknologiske insentiver (Reding, 2021). I møte med forbruker og kunde må man også tenke på hvordan de søker informasjon og hvordan din virksomhet kan komme de i forkjøpet.

**Bærekraft** og miljø er kanskje mer viktig for forbrukere og kunder nå enn noensinne. Samfunnsmessige trender som grønnere livsstil og økt press på å redusere hastigheten på den globale oppvarmingen påvirker næringslivet i stor grad (Reding, 2021). Bærekrafts faktorer vil være global oppvarming, karbonavtrykk, forsyningskjeder og produksjon.

**Juridiske** faktorer handler om hvilke lover og regler din virksomhet er pliktig til å følge. Dette både nasjonale og internasjonale reglement din virksomhet må forholde seg til, alt ettersom hvordan dere opererer. Her må man også ta for seg juridiske faktorer som kan føre til at du må endre prosedyrer eller arbeidsprosesser for å møte nye lovkrav (Reding, 2021). Denne overlapper til en viss grad med politiske faktorer. Forskjellen ligger i at juridiske faktorer diktere alltid hva en virksomhet lovbestemt kan eller ikke kan gjøre, og må derfor alltid overholdes.

## 2.13 Data-drevne makroanalyser

Data-drevne makroanalyser vil si å bruke data og datavitenskapelige teknikker til å analysere makroomgivelsener. I kontekst av denne oppgaven utgjør dette makroomgivelsene virksomheter står ovenfor, med andre ord PESTEL-faktorene. Basert på gjennomførte litteratursøk, er det lite eksisterende informasjon å hente om dette temaet. Likevel kan det antas at fordelene data-drevne analyser generelt utgjør, også vil være gjeldene ved bruk på data fra makroomgivelser. Bruk av data-drevne makroanalyser kan dermed potensielt gi virksomheter flere fordeler. Overordnet kan det gi virksomheter bedre beslutningsgrunnlag ved å levere innsikt i trender og mønstre i makroomgivelsene, som igjen kan hjelpe ledere med å ta bedre og mer informerte beslutninger. Ved å benytte data framfor intuisjon og subjektive meninger, kan data-drevne makroanalyser bidra til at virksomheter i større grad er i stand til å ta mer objektive beslutninger. Videre kan bruk av data-drevne makroanalyser automatisere prosesser som tidligere krevde manuell innsamling og analyse av data, noe som kan frigjøre tid for ansatte til å fokusere på andre oppgaver. Automatisering vil også gjøre det mulig for virksomheter å raskere bli bevisst på endringer i makroomgivelsene deres. Dette kan hjelpe virksomheter med å identifisere muligheter og trusler i omgivelsene, og

dermed hjelpe dem med å tilpasse seg raskere og mer effektivt. Bruk av data-drevne makroanalyser vil også kunne hjelpe virksomheter med å lage mer nøyaktige prognoser om makroomgivelsene, for eksempel med tanke på markedsutvikling, endringer i regelverk, økonomiske og politiske forhold, etc. Dette kan hjelpe virksomheter med å forberede seg på fremtidige endringer, og bedre planlegge for risiko og muligheter som oppstår.

Mange virksomheter har de siste årene tatt store steg for å bedre utnytte data samlet gjennom interne aktiviteter og prosesser. Sammenlignet har svært få gjort det samme for å realisere potensialet av å benytte data generert på utsiden av deres egen virksomhet, for å koble mot eller kontekstualisere internt generert data. Effektiv bruk av ekstern data for å analysere og gi innsikt i makroomgivelsene en virksomhet befinner seg i, har et stort potensial til å hjelpe virksomheter med å ta riktige beslutninger til riktig tid, og gjennom dette oppnå konkurransefortrinn (Aaser & McElhaney, 2021). En grunn til den synkende graden av rapporterte konkurransefordeler knyttet til bruk av dataanalyser i beslutninger, er at fler og fler virksomheter evner å ta det i bruk nok til å forholde seg konkurransedyktige. Likevel, sliter mange med å utnytte det på en måte som gir konkurransefordeler (Ransbotham et al., 2016). De konkurrerende aktørene blir likere, og konkurransen dermed utjevnet. Virksomhetene som ønsker å lede, må innovere. Data-drevne makroanalyser kan potensielt være en kilde til fordeler for mange virksomheter, dersom de evner å ta det i bruk.

## 2.14 Dataarkitektur

For at virksomheter effektivt skal være i stand å benytte seg av dataanalyser, må de ha verktøy som understøtter deres behov for data. En måte å gjøre dette på, er gjennom å utarbeide en god dataarkitektur. Dataarkitektur defineres i The Open Group Architecture Framework som «*en beskrivelse av strukturen og samspillet til virksomhetens viktigste datakilder, datatyper, logiske dataressurser, fysiske dataressurser og ressurser for datahåndtering*» (TOGAF, u.å.). Dataarkitekturen gir et overordnet perspektiv og veileder hvordan data samles inn, integreres, forbedres, lagres og leveres til ansatte som benytter det i arbeidet deres. Dette hjelper til ved å gjøre data tilgjengelig, nøyaktig og fullstendig, slik at det kan benyttes i forretningsbeslutninger (Sherman, 2015, s. 107–108). Når nye datakilder oppstår, sikrer en god dataarkitektur at dataen er håndterbar og nyttig. Mer spesifikt kan det unngå unødvendig datalagring, og forbedre datakvalitet gjennom rensing og fjerning av dupliserte data. Moderne dataarkitekturer gir også mekanismer for å integrere data på tvers av områder, for eksempel mellom avdelinger eller forretningsenheter. En moderne dataarkitektur gjør det også mulig å bryte ned datasiloer uten den enorme kompleksiteten som følger med å lagre alt på ett sted. (IBM, u.å.) Dataarkitektur deles ofte inn i lag (eng. layers) for å kategorisere og organisere komponentene som inngår. J.Y. Zhu et al. (2019) foreslår at virksomheter som håndterer stordata deler deres dataarkitektur inn i fem lag. Disse er innsamlingslag (eng. collection layer), lagringslag (eng. storage layer), databehandlingslag (eng. data processing layer), analyselag (eng. analytics layer), og applikasjonsslag (eng. application layer) (Zhu et al., 2019).

*Innsamlingslaget* fungerer som grunnlaget i dataarkitekturen (Zhu et al., 2019). Det er her selve dataen samles inn. Med stadig flere datakilder, trengs et innsamlingslag for å integrere både strukturert og ustrukturert flerkilde data for videre bruk. Dataen som samles inn vil bli ledet inn i databehandlingslaget, og den oppsamlede dataen vil bli lagret i lagringslaget. Deretter blir dataen analysert gjennom bruk av spesifikke

analyseverktøy i analyselaget, basert på etterspørselen fra applikasjonslaget. Løsninger for å hente data i dette laget kan blant annet være åpne datakilder, tradisjonelle søkemotorer, søkemotorer dedikert til datasett, tredjepartsleverandører som data-leverandører og meglere, data fra sosiale medier, eller data fra eksperter (Strand and Carlsson, 2008; Roeder et al., 2020; Strand and Syberfeldt, 2020; Krasikov et al., 2021; som sitert i Legner et al., 2022. s. 9). Åpne datakilder er data som utgis til offentligheten fra f.eks. staten eller organisasjoner, som virksomheter deretter kan hente inn gratis. I Norge har vi for eksempel plattformen data.norge. Denne plattformen består av innhold som blant annet datasett, levert av offentlige og private aktører. Man kan også få tilgang til publiserte API-er (eng. application programming interface). En API er et programvaregrensesnitt som gir direkte tilgang til data og funksjonalitet i et datasystem, som man enkelt kan hente et eller flere datasett gjennom (Rossen & Nätt, 2022). Norske virksomheter kan enkelt hente ut innhold fra data.norge (Digitaliseringsdirektoratet, 2022). Tredjepartsleverandører er aktører som samler inn data via en rekke ulike kilder, med mål om å videreselge dataen. Data fra tredjepartsleverandører vil ofte bli rensket og tilpasset bruksbehov før salg. Slik data vil dermed ofte være av høyere kvalitet (Zhu et al., 2019).

*Lagringslaget* fungerer som det andre laget og innbefatter lagringsinfrastruktur for datalagring (Zhu et al., 2019). Det er mange måter å lagre data på, og virksomheter burde velge dette avhengig av deres behov. Noen virksomheter benytter seg av flere ulike måter å lagre dataen deres på. Typiske måter å lagre data på er blant annet gjennom distribuerte datamiljø som Hadoop Distributed File System (HDFS). Dette er den mest brukte løsningen av sitt slag, ettersom HDFS er kostnadseffektivt, skalerbart og enkelt å ta i bruk (Zhu et al., 2019). En annen måte å lagre data på, er ved bruk av relasjonsdatabase-styringssystemer (RDBMS) (Mysore et al., 2013). Ulike metoder har ulike fordeler og ulemper, for eksempel med tanke på skalerbarhet. Dersom en virksomhet eksempelvis har kunder som genererer et relativt lavt antall transaksjoner per dag, vil tradisjonelle relasjonsdatabaser kunne tas i bruk med hensyn på responstiden. Hvis brukerne derimot genererer mer enn en million transaksjoner per dag, vil ikke relasjonsdatabasen kunne takle dette, og HDFS-lagringsystemer vil dermed være et bedre valg for å garantere god datagjennomstrømming og responstid (Zhu et al., 2019).

*Databehandlingslaget* er kjernen i dataarkitekturen (Zhu et al., 2019). Dette er et grunnleggende lag for mange dataanalyseoppgaver. Datakilder kan være ekstremt heterogene i struktur og innhold. Disse forskjellige datakildene har stor forskjell i datakvalitet, for eksempel med tanke på dekningsgrad og nøyaktighet. Generelt utfører databehandlingslaget parallell databehandling, datarensing, dataintegrasjon, datafusjon, dataindeksering, virtualisering med mer, for å gjøre dataen klar til analyse.

*Analyselaget* er laget hvor dataen blir analysert (Zhu et al., 2019). Det vil kreve ulike verktøy for dataanalyse, avhengig av brukerbehov. Eksempler på verktøy som kan benyttes her er Apache Spark, Power BI og Tableau. Programmeringsspråk som R og Python kan også benyttes, samt en rekke ulike tilleggsverktøy til alle disse utviklet for mer spesifikke behov. Ulike verktøy støtter ulike funksjoner. Tableau har for eksempel et større fokus på ren datavisualisering enn Spark, som har størst fokus på dataanalyse og maskinlæring (Bonthu & Bindu, 2017; Zhu et al., 2019).

*Applikasjonslaget* er det øverste laget (Zhu et al., 2019). Gjennom valg av metoder og verktøy i de fire underliggende lagene, kan virksomheter bygge applikasjoner for ulike

ansatte i virksomheten. Dette kan for eksempel være applikasjoner for forretningsinnsikt, markedsanalyse, eller prediksjoner.

## 2.15 Datavarehus og datasjø

Dersom data skal lagres over lengre tid og ikke bare for å støtte daglige operasjonelle aktiviteter, for eksempel til senere analytiske formål, krever dette systemer som støtter store mengder data (Panoply, u.å.; SAP, u.å.). Virksomheter lagrer ofte store mengder historisk data i det som kalles et datavarehus (eng. data warehouse). Et datavarehus er ethvert system som samler data fra en rekke kilder innen en virksomhet. Dette kan eksempelvis være ERP (Enterprise Resource Planning) og CRM (Customer Relationship Management)-systemer, ulike databaser, eller eksterne kilder som data fra partnere og IoT (Internet of Everything)-enheter. Datavarehus benyttes for å oppbevare og sentralisere data virksomheten besitter, slik at hele virksomheten kan benytte seg av den til ulike formål. Tradisjonelt sett er datavarehus lokalisert hos virksomheten selv, såkalt «on-premise». Virksomheten eier maskinvaren, og har egne ansatte for å drifte den. I nyere tid har derimot flere virksomheter valgt å heller benytte skybaserte datavarehus, da dette ofte har fordeler knyttet til blant annet kostnad og skalerbarhet (Panoply, u.å.; SAP, u.å.).

Rundt 2010 dukket også et nytt konsept opp kalt datasjø (eng. data lake), og siden den gang har mange virksomheter begynt å ta det i bruk for å lagre data (Miloslavskaya & Tolstoy, 2016). En datasjø refererer til en enormt skalerbar lagringsplass for data som kan lagre store mengder rådata i sitt opprinnelige format, fram til det hentes ut av datasjøen til videre bruk. I tillegg til dette, består en datasjø av behandlingssystemer som kan ta inn data uten at datastrukturen kompromitteres (Laskowski, 2016, som sitert i Miloslavskaya & Tolstoy, 2016, s. 302). Datasjøer blir vanligvis utviklet til å håndtere store og raskt ankomne mengder ustrukturerte data, noe som kan gi ytterligere innsikter ved bruk. Dette i motsetning til datavarehus, som vanligvis består av strukturert data. En datasjø skiller seg også ut fra datavarehus, ved at data som opprettes umiddelbart blir tilgjengelig for uthenting (Miloslavskaya & Tolstoy, 2016).

## 3 Metode

Kapittel 3 Metode skal ta for seg forskningstilnærmingen til prosjektet. Dette vil innbefatte en stegvis beskrivelse av prosessen, hvordan vi har gått frem, og hvilke metoder vi har brukt. Kapittelet ender i to separate resultatkapitler. Dette for å svare på de to ulike momentene problemstillingen vår tar opp.

### 3.1 Forskningstilnærming: Design Science Research Methodology

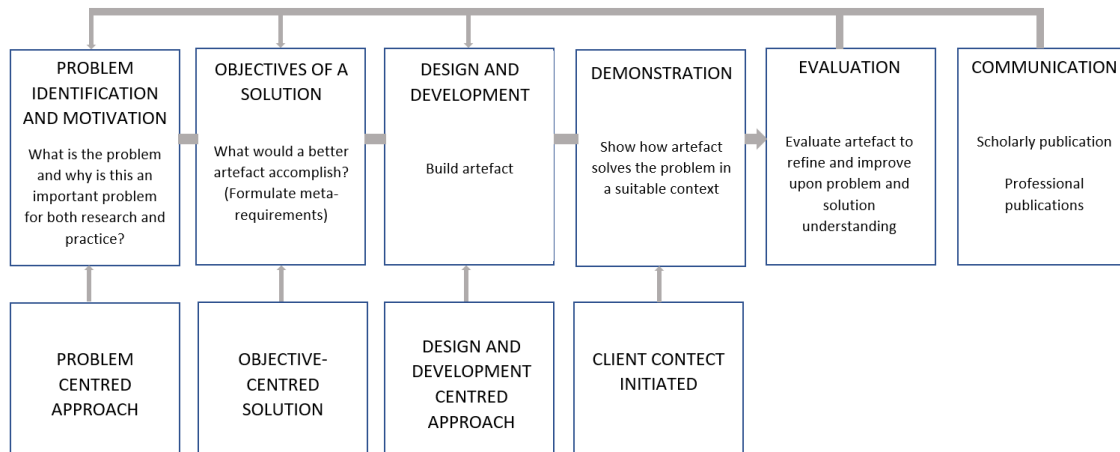
Gjennom formuleringen av problemstillingen vår ble det tidlig klart at vi måtte benytte oss av en kvalitativ forskningsmetode. Siden vi skulle utarbeide krav trengte vi informasjon om eksempelvis eksisterende arbeidsprosesser, kompetanse og data som virksomheter innehar, og som vi dermed kan hente ut og prosessere. Videre arbeid med problemstillingen viste at det ikke holdt å snevre det inn til kvalitativ metode, vi måtte vær enda mer spesifikk. Vi har valgt å tilnærme oss denne studien som et Design Science Research-prosjekt. DSR er egnet for å besvare oppgaven fra twoday Avento fordi hovedmålet med Design Science Research er å opparbeide seg kunnskap om fagområdet slik at man kan utvikle et artefakt eller løsning basert på dette. I tillegg kan forskningsmetoden fungere som en iterativ prosess der man kommer med et forslag, legger dette frem, får det evaluert, og går tilbake et steg dersom løsningen ikke er tilfredsstillende nok.

Design Science Research Methodology (DSRM) er et problemløsningsparadigme som ønsker å forbedre kunnskap ved å lage innovative artefakt som løser problemer og forbedrer omgivelsene der de er plassert. Metoden har en velutviklet tilnærming og beskrivelse, samtidig som den dekker sentrale aspekter ved utvikling av et rammeverk. Design Science (DS) har som formål å skape ting som tilfredsstillers menneskers behov (Peffer et al., 2014). DSRM holder på det samme formålet, i tillegg til at det er en forskningsmetode der man har en velutviklet, systematisk tilnærming til problemet. Basert på metodens tydelige konstruksjon, vil man effektivt vite hvordan vi skal gå fram, samt handle (Peffer et al., 2014). DSRM vil derfor bidra til at man dekker grunnleggende og avgjørende områder for å lykkes når rammeverket skal utvikles (Peffer et al., 2014). Stegene vil også forsikre at rammeverket går gjennom kritiske faser som fokuserer på selve designet og utviklingen, at det blir demonstrert, og avslutningsvis evaluert før utgivelse (Peffer et al., 2014).0020

Ved bruk av DSRM må man først klargjøre hva slags type problem man står overfor og derfor hvilken tilnærming man skal ha til metoden. Hvilken tilnærming man har, vil videre avgjøre hvor i prosessen man begynner. Man kan ha en problemsentrert, målsentrert, design og utviklingsentrert, eller en klientforbindelse-basert tilnærming. Dersom man begynner med en problemsentrert tilnærming, vil man begynne helt til venstre i modellen med å definere problemet. Deretter skal en løsning eller artefakt skisseres sammen med mål for løsningen. Basert på dette vil man videre begynne en design- og utviklingsprosess. Etter dette følger en demonstrasjon og evalueringsprosess med påfølgende tilbakeføring til et tidligere steg for å videreutvikle løsningen basert på



**Figur 1 Design Science Research Methodology. Modellen er laget basert på Peffers modell for DSRM (Peffers et al., 2014)**

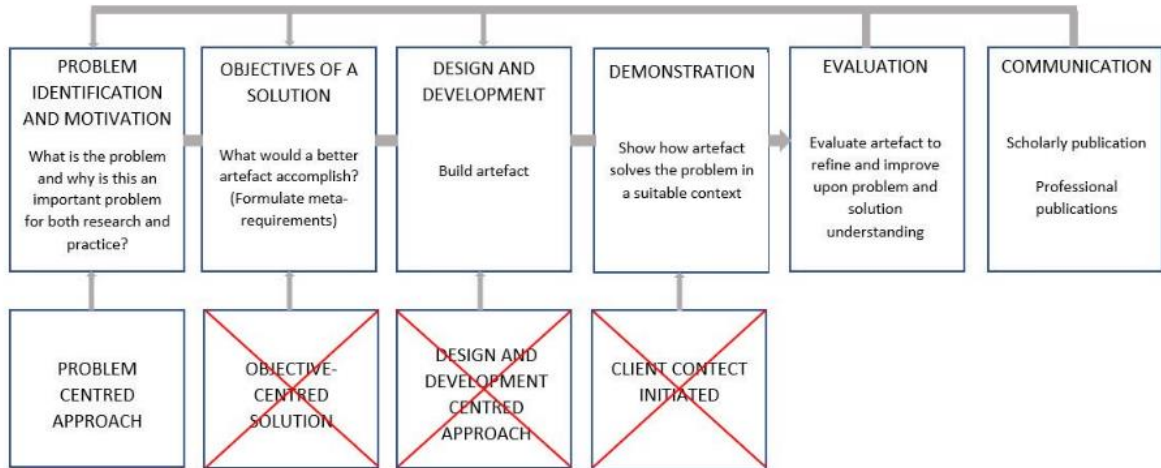


læring fra evalueringen, dersom dette er nødvendig. Dette gjøres frem til et tilfredsstillende resultat er oppnådd og artefakt eller løsning kan kommuniseres. Vi vil i resten av kapittelet gjøre rede for hvordan vi gjennomførte studien som et design science research-prosjekt.

### 3.2 Steg 1 Problemidentifisering og motivasjon

Denne studien er basert på en løselig formulert problemstilling fra oppdragsgiveren twoday Avento. Oppgaven hadde som hensikt å «...gi en innsikt i hvordan digitalisering kan hjelpe organisasjoner i de utfordrende situasjonene vi ser i verdensbildet i dag.» Etter en innledende dialog med oppdragsgiver bestemte vi oss for å tilnærme oss oppgaven med Design Science Research, og med mål om å utforme et rammeverk som twoday Avento kunne bruke for å tilby utvikling av data-drevne makroanalyser til kundene sine. Fordi oppgaven var stilt såpass åpen, uten en helt klar problemformulering, valgte vi å initiere DSRM-prosessen med en problem-sentrert tilnærming.

Det første steget i en problem-sentrert tilnærming er å identifisere problemet og deretter begrunne verdien av en løsning. Videre skal motivasjonen bak løsningen begrunnes. Problemdefinisjonen vil gjennomgående bli brukt til å utvikle en løsning. Det er derfor viktig å være så spesifikk som mulig. Dette fordi man ønsker å utvikle en løsning som kan dekke problemets kompleksitet. Ved å godt begrunne verdien av løsningen, vil man oppnå to ting; en større sannsynlighet for at utenforstående forstår problemet, i tillegg til en større sannsynlighet for aksept av løsningen. Det hjelper også å forstå drøftingen knyttet til utviklerens forståelse av problemet og viktigheten av løsningen.



**Figur 2 Vår tilnærming til DSRM**

For å utarbeide en bedre problemforståelse gjennomførte vi et litteratursøk for å utforske hvilket bredere forskningstema oppgaven berører. Først søkte vi å utforske hvilket bredere forskningstema oppgaven berører. Vi konkluderte med at oppgaven handler om data-drevne beslutningsprosesser, og da spesifikt strategiske beslutningsprosesser knyttet til data-drevne makro-analyser av omgivelsene. Vi valgte å tilnærme oss oppgaven som en form for datadreven beslutningsprosess fordi oppgavestiller framhevet bruk av data-analyser som måte å håndtere de ulike utfordringene virksomheter støter på i verdensbildet i dag. Med dette som utgangspunkt i utviklingen av problemstillingen, tok gjorde vi et videre litteratursøk om hvordan virksomheter i dag bruker makroanalyser. Noe som stadig var en gjenganger i litteraturen, var bruk av data for å forstå makroomgivelsene virksomheter befant seg i. Gjennom studiet vårt har vi flere ganger vært innom temaer knyttet til kompleksiteten av å gjennomføre digitaliseringsprosjekter, både med tanke på tekniske faktorer, men også organisatoriske. Når det kommer til bruk av data for å forstå makroomgivelsene til virksomheter, var dette noe vi ønsket å undersøke videre.

Videre snevret vi inn problemstillingen, ved å fokusere på litteratur knyttet til virksomheters bruk av makroanalyser. Basert på flere kilder i litteraturen, var konklusjonen vår at virksomheter flest forstår viktigheten som ligger i bruk av analyser i forretningsbeslutninger, men at svært få klarer å gjennomføre de omstillingene som kreves for å benytte seg av det. Når det kommer til bruk av ekstern data i analyser av makroomgivelser, genererer ikke virksomheter all data som kreves internt. Potensialet rundt bruk av ekstern data i analyser i virksomheter er dermed stort. Rapporten til Berndtsson et al. (2020), identifiserte et behov for å utvikle endringsstrategier som tok for seg både tekniske og ikke-tekniske (organisatoriske) faktorer samtidig.

Med utgangspunkt i dette litteratursøket, så vi derfor at det fantes få praktiske råd knyttet til data-drevne makroanalyser. Vi så derfor behovet for et rammeverk som kan bistå virksomheter i å gjennomføre slike analyser på en hensiktsmessig måte. Vi stilte derfor følgende problemstilling: «Vi vil finne ut hva som kreves teknisk og organisatorisk av et rammeverk for utvikling av data-drevne makroanalyser. Med bakgrunn i dette ønsker vi å forsøke å skissere et rammeverk.». Før vi kunne utvikle rammeverket, måtte vi først formulere mål og krav til løsningen.

### 3.3 Steg 2 Formulere mål for og krav til løsningen

Det andre steget i DSRM er å formulere mål og krav til løsningen som skal utvikles. I denne fasen skal man skissere rammene for løsningen basert på problemdefinisjonen, og eksisterende kunnskap om hva som er mulig og gjennomførbart. Det er viktig her å begrunne hvordan denne nye løsningen forventes å løse problemer som hittil ikke er behandlet. Hensikten bak løsningen må utledes fra problemdefinisjonen. For å gjøre dette må man ha kunnskap om problemets tilstand og den aktuelle løsningen, hvis noen, og løsningenes effektivitet.

Etter å ha konstatert at det var mangel på endringsstrategier som tok for seg tekniske og organisatoriske faktorer samtidig, samt identifisert behovet for et slik rammeverk, begynte vi å definere rammeverket, hva rammeverket skulle og ikke skulle ta for seg. Et konseptuelt rammeverk inkluderer en eller flere formelle teorier samt andre begreper og empiriske funn fra litteratur (Lemieux, u.å.). Et konseptuelt rammeverk brukes til å vise sammenhenger mellom disse ideene og hvordan de relaterer seg til forskningsstudiet. Konseptuelle rammeverk sees ofte i kvalitativ forskning ofte fordi en teori ikke fullt ut kan adressere fenomenene som studeres. Videre anså vi det som optimalt å bygge rammeverket vårt på prinsipp fra prosjektledelses- og gevinststyringsrammeverk.

Et *prosjektledelsesrammeverk* (Eng. Project management framework) er en overordnet metode for gjennomførelse av et prosjekt. Rammeverket er konstruert for å oppnå sine prosjektmål innenfor begrensninger en virksomhet har satt. Prosjektledelsesrammeverk vil gjøre dette gjennom retningslinjer, praksiser for prosjektledelse, og prosedyrer. Det foreligger også fokusområder som fokuserer på å oppnå bedre resultater, effektiv ressursbruk, og økt tilfredshet hos kundegruppene. Et kjente eksempel på prosjektledelsesrammeverk er PRINCE2 (Badewi, 2022).

Et *Gevinststyringsrammeverk* (Eng. Benefits management framework) har som formål å realisere fordeler og gevinster som er forventet for en virksomhet. Rammeverket er en sikkerhet for at virksomheter skal lykkes med å oppnå de vedtatte strategiene og målene sine (Badewi, 2022). Gjennom bruken av rammeverket vil man kontinuerlig ha et overblikk over gevinster som realiseres gjennom prosjektets livssyklus.

Rammeverket vårt er ment å gi en oversikt over tekniske og organisatoriske faktorer som bør/må ligge til grunn for at en virksomhet skal være i stand til å utvikle data-drevne makroanalyser, og at implementeringen av disse skal gå så smidig som mulig. Rammeverket skal ikke ta for seg virksomheten i seg selv. Dette vil si at vi ikke tar høyde for virksomhetens størrelse, deres tilgjengelige ressurser, eller hvilket marked eller næring de driver innenfor. Rammeverket skal heller ikke ta for seg en detaljert forklaring over hvilken teknologi som må ligge til grunn. En grunnleggende beskrivelse vil innbefattes i rammeverket. Dette fordi det kommer an på hvilken bedrift og av hvilken størrelse og marked.

#### 3.3.1 Kravhåndtering

For å bedre være i stand til å forstå kravene mulige brukere vil stille til løsningen, starter vi prosessen med å utarbeide krav til løsningen. Under et hvert utviklingsprosjekt, burde teamet forsøke å besvare tre fundamentale spørsmål (Pohl & Rupp, 2015, som sitert i Volk et al., 2017, s. 4): *Hva er kravhåndtering (eng. requirements engineering)? Hva er kravene til produktet? Hvem er prosjektets interessenter, og hvilken rolle spiller de?* For å besvare det første fundamentale spørsmålet til Pohl & Rupp, baserer vi oss på definisjonen til Zhi Jin (2018). Jin forklarer at kravhåndtering er hele prosessen med å få

fram behov og ønsker fra interessenter, og deretter etablere et sett med avtalte, fastsatte og detaljerte krav som vil fungere som et grunnlag for videre utviklingsaktiviteter. (Jin, 2018). I prosessen med å utvikle rammeverket vårt, vil det være viktig å forstå kravene som mulige interessenter stiller til produktet. Gjennom å sikre at kravene møtes, øker vi verdien av løsningen ovenfor brukere, og samtidig sannsynligheten for at løsningen faktisk blir tatt i bruk. Krav til løsningen vår, kan defineres som «*en dokumentert representasjon av en betingelse eller evne som trengs av en bruker for å løse et problem eller oppnå et mål*» (IEEE, som sitert i Wahono, 2003).

Vi stadfestet tidlig at vi ønsket å benytte en kombinasjon av anskaffelse og fremkalling av krav videre i kravhåndteringsprosessen. De to begrepene brukes ofte litt om hverandre. Generelt kan man si at anskaffelse handler om å *hente inn krav som allerede er synlige* (f.eks. gjennom eksisterende dokumenter), og at fremkalling handler om å *synliggjøre krav som ikke enda er dokumentert* (f.eks. gjennom intervjuer) (AssistKD, 2021). Basert på det faktum at det allerede eksisterte noe litteratur rundt mulige tekniske og organisatoriske behov virksomheter har, dersom de ønsker å benytte seg av data-drevne makroanalyser, valgte vi å benytte prinsipper fra kravanskaffelse her. Tatt i betraktning av at vi har mulige informanter tilgjengelig i twoday Avento, ønsket vi å benytte denne muligheten i en tilnærming til fremkalling av krav. Til tross for at vi ikke utviklet rammeverket spesifikt for deres virksomhet, betraktet vi det fortsatt som svært nyttig å anse twoday Avento som en representant for virksomheter som tilbyr slike tjenester. De har i tillegg ansatte som besitter relevant kunnskap og erfaring innenfor områder rammeverket vårt må adressere. Vi anså det dermed som verdifullt å gjennomføre kvalitative intervjuer med ansatte fra ulike deler av virksomheten, med det formål om å benytte svarene som en veiledning, samt bekreftelse på funnene våre. Rammeverket vårt må adressere et stort antall ulike tekniske og organisatoriske faktorer, så svarene vi får fra informantene ville tenkelig fungere som en nyttig styrende ressurs videre i arbeidet vårt med utviklingen av rammeverket.

Tatt i betraktning av at vi ønsker å utvikle et rammeverk som ikke er spesifisert til bruk av én bestemt virksomhet, så vi for oss at prinsipper fra Market Driven Requirements Engineering (MDRE) ville være særlig aktuelt for oss. Denne fremgangsmåten tar for seg utvikling av produkter rettet mot større markeder. I en kravhåndteringsprosess, er det avgjørende å være bevisst på forskjeller i prosessen med kundespesifikk utvikling kontra markedsdrevet utvikling (Aurum & Wohlin, 2005, s. 95). Basert på litteratur om kravhåndtering innen program- og maskinvareprosjekter, samt prinsipper innen MDRE, valgte vi å arbeide med en definert kravhåndteringsprosess tilpasset vår type produkt, samt markedet vi utviklet løsningen for.

Stegene vi benyttet i kravhåndteringsprosessen ville dermed være:

1. kravinnhenting (fremkalling og anskaffelse),
2. kravanalyse,
3. kravdokumentasjon,
4. krav- verifisering og validering.

Når dette var gjennomført, kunne vi starte arbeidet med å designe selve løsningen.

### 3.3.2 Interessentanalyse

I prosessen med å utarbeide krav stod vi ovenfor en rekke ulike potensielle nøkkelinteressenter vi kunne intervju. For å svare på det tredje fundamentale spørsmålet til Pohl & Rupp (2015, som sitert i Volk et al. s. 4); «*Hvem er prosjektets interessenter, og hvilken rolle spiller de?*», samt bedre forstå hvordan vi mest hensiktsmessig burde prioritere krav med hensyn på løsningens brukerverdi, gjennomførte vi en interessentanalyse. Vi identifiserte både interne og eksterne nøkkelinteressenter. De interne var toppledelse, dataanalytikere, IT-ansvarlige, avdelingsledere og øvrige ansatte. De eksterne var eksterne konsulenter, kunder eller klienter, myndigheter og investorer.

Interessentanalyse er oppsummert prosessen med å systematisk samle og analysere kvalitativ informasjon for å vurdere hvilke påvirkede aktørers interesser som burde tas i betraktning (Schmeer, 2000). I vårt tilfelle, ble dette med hensyn på rammeverkets funksjon. Interessenter er aktører, dvs. individer eller virksomheter med egeninteresse i det som blir utviklet. I prosessen med å gjennomføre en interessentanalyse, er det vanlig å vurdere de ulike interessentene etter makt, samt interesse i prosjektet (Smith, 2000). Produktet vårt er ikke ment for én spesifikk virksomhet, men heller et bredt spekter av virksomheter med felles mål om å utvikle bruken av makroanalyser enten internt i egen, eller eksternt i andre virksomheter. Twoday Avento er som tidligere beskrevet et konsultantselskap med spisskompetanse innen ulike digitale tjenester. Da vi arbeidet med å hente ut krav for å utvikle et rammeverk ment for et bredere marked av virksomheter, ville det være lite gunstig å kun betrakte interessenter med spisskompetanse innen fagområdet. Basert på dette så benyttet vi twoday Avento som en representant, og deretter valgte vi å trekke inn teoretiske interessenter som virksomheter tradisjonelt består av.

Ettersom rammeverket ikke utvikles for én spesifikk virksomhet, måtte vi også justere litt på bruken av begrepene makt og interesse, og tilpasse det konteksten av prosjektet vårt. Begrepet *makt* benyttet vi i den forstand av «*hvor stor rolle denne interessenten vil utgjøre i en prosess med utvikling i bruken av makroanalyser internt i en virksomhet, og hvor stork makt denne interessenten da teoretisk vil ha*». Ettersom utviklingen av rammeverket foregikk uavhengig av en spesifikk casebedrift, hadde vi ingen interessenter med direkte makt. Av denne grunn så vi det som mer nyttig å bruke begrepet makt slik definert over. *Interesse* kunne i denne konteksten forstås mer tradisjonelt som «*graden av interesse den aktuelle interessenten har i prosjektet*». I interessentanalysen ønsket vi å undersøke både interne og eksterne interessenter. Under interne interessenter, undersøkte vi de aktører som befinner seg internt i en typisk virksomhet, og som vil bli påvirket dersom rammeverket vårt blir tatt i bruk. Under eksterne interessenter derimot, undersøkte vi også de som blir påvirket, men som befinner seg utenfor en typisk virksomhet. Interessentanalysen vår resulterte i en forståelse av at toppledelse og avdelingsledelse, dataanalytikere, IT-ansvarlige og eksterne konsulenter, var de mest hensiktsmessige interessentene å ta med oss videre i prosessen med intervju av ansatte hos twoday Avento.

### 3.3.3 Kravinnhenting

Etter å ha identifisert interessentene vi ønsket å intervju, samt dannet en felles forståelse rundt hvilken informasjon vi ønsket å hente fra de, var vi bedre rustet til å starte prosessen rundt informasjonsinnhenting for etablering av krav til løsningen. For at vi skulle være i stand til å benytte kunnskap informanter i twoday Avento innehar som

kilde, begynte vi tidlig å diskutere med bedriftskontakt hvordan vi kunne gjennomføre intervjuer av deres ansatte. Videre fulgte utarbeidelsen av spørsmål til de ulike nøkkelinteressentene. Med hensyn til de utvalgte interessentenes arbeid, valgte vi å utvikle forskjellige sett med spørsmål til de ulike interessentene. Dette gjorde vi da ansvarsområdene deres krever ulik kunnskap og ferdigheter, og informasjonen vi trengte fra informantene varierte.

Vi utforsket mye litteratur rundt kravhåndtering, og undersøkte hvordan vi på best mulig måte kunne utforme spørsmål hvor de eventuelle svarene ville avdekke krav til rammeverket vårt. Det aller meste av litteratur rundt kravhåndtering, omhandlet utvikling av rene maskin- og programvareprodukter. Vi måtte derfor gjøre store tilpasninger i metodene vi benyttet for å utarbeide intervju spørsmål og måten vi fremkalte krav. Vi endte opp med å utvikle unike sett med spørsmål til eksterne konsulenter, IT-ansvarlige, dataanalytikere, og ledelse. Toppledelse og avdelingsledere ble slått sammen til ett sett med spørsmål, da informasjonen vi ønsket å anskaffe gjennom disse i stor grad var tilsvarende lik. Vi satte alle spørsmålene inn i et felles dokument, som også inkluderte en forklaring av oppgaven og artefaktet vi ønsket å skissere. Dette ba vi informantene lese i forkant av intervju spørsmålene, for å sikre en bedre forståelse av hva vi var ute etter av informasjon fra dem.

**Tabell 1 Intervjuobjekt**

|          | <b>Stilling</b>                                     | <b>Beskrivelse</b>  | <b>Fokus for intervjuet</b>   |
|----------|---|---------------------|---|
| <b>1</b> | Ansatt som seniorrådgiver                           | Intervju med ansatt | Få et innblikk i hvordan twoday Avento går frem og hva ser de etter i virksomheter i innledende samtaler. Hva trengs av data, hva er typisk ønske eller motivasjonen til virksomheten og hvem er pådragsyter. |
| <b>2</b> | Nylig ansatt i Business Intelligence avdelingen     | Intervju med ansatt | Fokus data, datakvalitet, ansvar for data. Vi spør også om typisk motivasjon bak å ønske å bli data-drevet, og hvilke trender de ser.   |
| <b>3</b> | Avdelingsleder for Business Intelligence avdelingen | Intervju med ansatt | Enda mer teknisk fokus. Ønsker å få informasjon om fra hvor og hvordan data samles inn. Spør også her om hva som er typisk motivasjon bak å bli data-drevet.  |
| <b>4</b> | Ansatt som ICT Operations Manager                   | Intervju med ansatt | Ønsket informasjon om samarbeid mellom IT- og BI-avdelingen. Vi ønsket også informasjon om IT-arkitektur, plattformer og data brokers.  |

Ved bruk av uthenting og anskaffelse av krav, satt vi igjen med informasjon fra primærkilder (intervju) samt sekundærkilder (søk i eksisterende litteratur). Basert på informasjonen vi hentet fra disse kildene, var vi i stand til å starte prosessen med å skissere krav for påbegynnelse av neste fase i kravhåndteringsprosessen.

### 3.3.4 Kravanalyse

Neste steg i arbeidet med kravhåndtering var å analysere kravene. Prosessen med kravanalyse handler om å analysere, validere og tilpasse kravene hentet ut i foregående faser (Visure, u.å.). I henhold til MDRE ønsket vi å gjøre dette gjennom en innledende

evaluering, spesifisering av krav i detalj, og deretter en utvelgelse av kravene som skulle realiseres. Disse kravene ville vi deretter dokumenteres på en oversiktlig måte. Dette gjorde vi for å sikre oss en oversikt over alle kravene som må adresseres i aktivitet 3 av DSRM, samt for å gjennomføre neste steg i kravhåndteringsprosessen.

Innledningsvis i dette steget, strukturerte vi kravene slik at de var enklere å håndtere og forstå. Videre gikk vi gjennom kravene, slik at vi kunne identifisere eventuelle motstridende krav, eller mangler. Avslutningsvis i dette steget, definerte vi kravene som var viktigst, og deretter de som var mindre viktige for rammeverket. For å forsikre oss om at kravene faktisk kunne gjennomføres, foretok vi oss en evaluering om hvorvidt kravene var forsvarlige, altså at de var realistiske og oppnåelige i henhold til en generell virksomhets ressurs og tidsplan.

### 3.3.5 Kravdokumentasjon

Kravdokumentasjon inngår som det tredje steget i kravhåndteringen. Hovedmålet ved dette steget er å dokumentere kravene på en tydelig og presis måte, slik at de kan forstås og implementeres, samt benyttes i neste steg av kravhåndteringsprosessen. Dermed startet vi med å skrive kravene i detalj. Her presiserte vi informasjon om kravenes kilde og beskrivelse. Deretter sørget vi for at kravene sto i stil med våre aktuelle interesser, før de ble vurdert og eventuelt tatt med videre. Dette inkluderer å sjekke om kravene samsvarer med virksomhetens nøkkelinteresser og deres kunnskap og andre ressurser. I et langsiktig perspektiv vil vi være nødt til å holde en oversikt over endringer i kravene, samt sørge for at alle versjoner er tilgjengelige for interessentene våre.

### 3.3.6 Krav-verifisering og validering

Det avsluttende steget i kravhåndteringsprosessen var å verifisere og validere kravene. Her var det viktig å skille mellom de to begrepene. Å verifisere kravene handler om å sikre at kravene ble dokumentert korrekt. Her måtte vi besvare spørsmålet; «skrev vi ned de riktige kravene?». Validering derimot, handler om å sikre at kravene gjenspeiler det faktiske behovet. Her måtte vi besvare spørsmålet; «er det dette som faktisk trengs?» (Terry Bahill & Henderson, 2005). Dette ble gjort iterativt gjennom samtaler med bedriftskontakt hos twoday Avento.

For å konkludere, valgte vi å utforme, samt gjennomføre, en kravhåndtering bestående av fire steg. Blant disse fire stegene finner vi en rekke aktiviteter som vi foretok oss, slik at vi forsikret oss om at gjennomførelsen av kravhåndteringen var suksessfull. Videre vil vi se på steg tre i DSRM, design og utvikling.

## 3.4 Steg 3 Design og utvikling

I dette steget skal løsningen designes. Forslag på slike løsninger kan være modeller, metoder eller konsept/teorier. Her skal også løsningens ønskede funksjonalitet og dens arkitektur bestemmes. Deretter skal man lage den faktiske løsningen. For å gå fra mål, til design og utvikling kreves ressurser som kunnskap om teori som kan tas med i en løsning.

I henhold til problemstillingen vår, ønsket vi å utvikle en struktur på rammeverket der brukeren enkelt får en oversikt over de ulike faktorene som må tas hensyn til. Dette er avhengig av hvor de befinner seg i prosessen med å utvikle data-drevne makroanalyser. Da prosessen med å utvikle produktet begynte, gjorde vi tidlig opp en tanke om at det

ville være hensiktsmessig å basere rammeverket på en inndeling av de elementer en virksomhet typisk består av. Det er mange årsaker til dette. Vi vet at virksomheter kan være svært utfordrende å forstå seg på, og at kompleksiteten påvirkes av faktorer som blant annet struktur og størrelse. Henry Mintzberg forklarte at «*en virksomhets struktur kan enkelt forklares som resultatet av måtene den deler sin arbeidskraft på i forskjellige oppgaver, og deretter oppnår koordinering mellom dem*» (Mintzberg, 1979). Nøyaktig hvordan dette gjøres, varierer derimot i stor grad. Alle virksomheter har unike strukturer. Dette er noe vi måtte ta høyde for i designet av rammeverket vårt, ettersom vi ikke ønsket at dette skal være et begrensende aspekt for å ta i bruk løsningen vår. Grunnet dette, ønsket vi å benytte et sett med felles dimensjoner alle virksomheter består av.

Tidlig i prosjektet formulerte vi at vårt prosjekt omhandlet datavitenskap, derav data-drevne metoder. For å underbygge disse fagfeltene, valgte vi å benytte oss av faglige ressurser gjennom primær- og sekundærkilder. I prosjektets tilfelle, baserte vi de faglige ressursene på intervjuer og litteratursøk. For å danne et grunnlag til førsteutkastet av kravene til rammeverket, valgte vi å bruke en kombinasjon av disse kildene, da de er i stand til å utfylle hverandre ytterligere, og dermed skape et mer komplekst bilde av kravene. Dermed besatt vi enkelte kilder man kunne basere kravene på, og deretter validere de gjennom ytterligere kilder. Samtidig vil man også være i stand til å tilføye de aktuelle kildene en ny dimensjon av informasjon, gjennom å bruke en kombinasjon av primær- og sekundærkilder. Som en ettervirkning, vil en kombinasjon av primær- og sekundærkilder bidra til en mer kompleks, legitim, og ansvarlig bruk av kilder, og dermed underbygge løsning i enda større grad.

For å skille de typiske fasene i en dataanalyseprosess, undersøkte vi flere ulike prosessmodeller for datavitenskapprosjekter. Vi undersøkte primært tre av de mest populære prosessmodellene; CRISP-DM, SEMMA (Sample, Explore, Modify, Model and Assess) og KDD (Knowledge Discovery in Databases) (Piatetsky, 2014). Valget falt til slutt på CRISP-DM. Dette begrunner vi i at modellen gir en strukturert tilnærming til alle typer datavitenskapprosjekter, at det med god margin er den mest brukte prosessmodellen, at den er fleksibel, og at den både har et større forretningsmessig fokus enn alternativene vi undersøkte.

Tatt i betraktning at vi ønsket et rammeverk som foretok seg organisatoriske og tekniske krav, og samtidig hadde hovedfokus på data-vitenskap og data-drevne makroanalyser, gjennomgikk vi diverse alternativer. For å finne alternativer å evaluere, henvendte vi oss til informasjonssøking gjennom både diverse litteratursøk og intervjuer. Blant forberedelsene som ble gjort, ble CRISP-DM og Buitelaars (2018) åtte dimensjoner synliggjort. Ved å ta utgangspunkt i en kombinasjon av CRISP-DM og Buitelaars åtte dimensjoner, fant vi ut at rammeverket ville bære tydelige preg av både tekniske og organisatoriske krav. CRISP-DM er som kjent en standardprosess for planlegging, utvikling og implementering av data-drevne prosjekter, og gir en strukturert tilnærming til datavitenskapprosjekter (IBM, 2011).

Denne tilnærmingen baserer på standardprosessens seks steg, forretningsforståelse, dataforståelse, dataforberedelse, modellering, evaluering og implementering. Mens CRISP-DM gir en strukturert tilnærming til datavitenskaplige prosjekter, gir Buitelaars åtte dimensjoner en vurdering av virksomhetens digitale modenhet i henhold til å være data-dreven. Dermed vil kombinasjonen av CRISP-DM seks steg og Buitelaars åtte dimensjoner komplimentere hverandre, slik at man tilegner seg en dypere forståelse av



hva CRISP-DMs steg består av. De åtte dimensjonene består av data, metrikker, ferdigheter, teknologi, ledelse, kultur, strategi og smidighet. De nevnte dimensjonene inneholder aspekter samtlige virksomheter baseres på, noe som vil være et fundament i vårt rammeverk. Dette gjør at vi dekker både de organisatoriske faktorene, samt de tekniske, gjennom denne kombinasjonen.

Ut ifra de åtte dimensjonene forklart av Buitelaar (2020, s. 33), legger vi et grunnlag for å adressere alle aspekter samtlige virksomheter består av. Dette er et viktig fundament i rammeverket vårt. Videre hadde vi et behov for å identifisere de ulike stegene som inngår i en makroanalyse. Gjennom dette var tanken og strukturert kunne gå fram og adressere de ulike tekniske og organisatoriske faktorene som gjør opp en virksomhet, samt hva som kreves for å utvikle makroanalyser i henhold til hvor de befinner seg i en makroanalyseprosess. Vi ønsket på bakgrunn av dette å danne en matrisemodell med de åtte dimensjonene identifisert av Buitelaar (2018) på den ene aksen, og en inndeling av stegene i en typisk dataanalyseprosess på den andre aksen.

**Tabell 2 Første forslag til utforming av rammeverk**

|                    | Forretningsforståelse | Dataforståelse | Dataforberedelse | Modellering | Evaluering | Implementering |
|--------------------|-----------------------|----------------|------------------|-------------|------------|----------------|
| <b>Data</b>        |                       |                |                  |             |            |                |
| <b>Metrikker</b>   |                       |                |                  |             |            |                |
| <b>Ferdigheter</b> |                       |                |                  |             |            |                |
| <b>Teknologi</b>   |                       |                |                  |             |            |                |
| <b>Ledelse</b>     |                       |                |                  |             |            |                |
| <b>Kultur</b>      |                       |                |                  |             |            |                |
| <b>Strategi</b>    |                       |                |                  |             |            |                |
| <b>Smidighet</b>   |                       |                |                  |             |            |                |

Tanken bak dette er å etablere grunnlaget for å systematisk gå fram for å adressere de ulike tekniske og organisatoriske faktorene som kan legge til rette for, eller som eventuelt kan hindre virksomheter i prosessen med å utvikle data-drevne makroanalyser. Vi ønsket at matrisefaktorene skulle danne et fundament som gjør det mulig å adressere alle potensielle faktorer på en systematisk og ryddig måte. I henhold til problemstillingen, ville dette også bidra til en oversiktlig løsning som vil være relevant og forståelig for alle virksomheter som ønsker å utvikle data-drevne makroanalyser.

For å oppsummere, ønsker vi å bruke en kombinasjon av CRISP-DMs seks steg, og Buitelaars åtte dimensjoner til å utforme løsningen vår. Et rammeverk bestående av denne kombinasjonen, vil tilføre en bedre forståelse av aktuelle aspekter en virksomhet enten består av, eller bør bestå av. Denne kombinasjonen gir dermed et dypere innsyn i CRISP-DMs seks steg, noe som gjør det lettere for utenforstående å forstå betingelsene for de tekniske og organisatoriske faktorene i en virksomhet. Videre vil vi ta for oss steg fire i forskningstilnærmingen vår, altså demonstrasjon.

### 3.5 Steg 4 Demonstrasjon

Her skal løsningen demonstreres for å løse en eller flere tilfeller av problemet. Dette kan innebære bruk av eksperimentering, simulering, case, bevis eller en annen passende aktivitet. Det er også viktig at de som demonstrerer innehar effektiv kunnskap om hvordan man bruker løsningen.

Etter å ha gjennomført en kravhåndteringsprosess, samt designet og utviklet kravene til rammeverket, skulle resultatet sendes til oppgavestiller for gjennomlesning og evaluering. Kravene ble lagt inn i et eget dokument, sammen med en forklaring av hva vi ønsket å utvikle. Dette for å sikre at de som leste gjennom forsto konteksten rundt kravene. Basert på dette fikk vi tilbakemelding som vil bli presentert i evalueringen.

### 3.6 Steg 5 Evaluering

I evalueringen skal man observere og måle hvor godt løsningen løser problemet. Dette innebærer å sammenligne målene med faktiske observerte resultater fra bruken av løsningen i demonstrasjonene. Derfor kreves det også kunnskap om relevante beregninger og analyseteknikker, og hvilke som er relevante avhenger av arten til problemet og løsningen. Teknikkene som tas i bruk kan sammenligne løsningsmålene fra aktivitet 2, kvantitative mål for systemutnyttelse f.eks. responstid eller tilgjengelighet. Konseptuelt kan en slik evaluering inkludere alle passende empiriske og logiske bevis.

Tilbakemeldingene var oppsummert at vi ikke hadde vært konkrete nok i formuleringen av kravene. Ønsket var at vi skulle være mer konkrete med hvilke aktiviteter virksomheter kunne gjennomføre i fasene av modellen vår, i tillegg til hva man skulle sitte igjen med etter at fasen var gjennomført. Som en prosessbeskrivelse tar CRISP-DM mange momenter ved implementeringen av datavitenskapprosjekter. Likevel er dette ikke nødvendigvis like tilpasset opp mot data-drevne makroanalyser. Dette gjelder hovedsakelig i de tre første fasene av modellen; forretningsforståelse, dataforståelse og dataforberedelse.

Med bakgrunn i denne konklusjonen, valgte vi etter dette å hovedsakelig ta et dypdykk i disse tre første fasene i modellen. De tre siste (modellering, evaluering og implementering) skulle likevel utgreies og forklares. Dette fordi vi erfarte at disse tre ikke krever like mye tilpasning spesifikt opp mot data-drevne makroanalyser. De tre siste fasene bærer preg av mer standardiserte prosesser og fremgangsmåter for å oppnå ønskede resultater. Data-drevne makroanalyser vil kreve et mer detaljert og komplekst syn på forretnings- og dataforståelse. Dette fordi man gjerne vil finne ut mer om virksomhetens plassering i marked og hvilken tilnærming man må ha til problemstillingen. To typiske tilnærminger til dette kan være om man ønsker å se etter gevinstene data-drevne makroanalyser kan tilføre bedriften eller om man ta utgangspunkt i dataene virksomheten innehar og ser etter nye måter å utnytte disse på. Basert på disse innvendingene, ville prosessen videre være med hovedvekt på virksomheten, dens plassering i et marked, dataene de innehar, hva de er i stand til å skaffe og bearbeiding av disse.

Rammeverket var originalt tenkt å se ut som en matrisetabell som vist i tabell 2. Dette resulterte i at overgangen fra fase 1 forretningsforståelse, til fase 2 dataforståelse ble vanskelig å konkretisere. Dette gjorde også slik at de åtte dimensjonene vi hadde definert, satte til en viss grad begrensninger for hva virksomhetene kunne undersøke i analysene. Samtidig syns vi at utseendet på rammeverket kunne bidra til at virksomheter låste seg fast til å kun se etter de åtte dimensjonene når de skulle gjennomføre analysene. Med bakgrunn i dette valgte vi å flytte forretningsforståelsesfasen til å ligge før de åtte dimensjonene er definert.

I tillegg konkluderte vi med at forretningsforståelsesfasen ikke inneholdt nok analyse til å kunne begrunne det fasen er ment å oppnå. Vi valgte å legge til Benefit Dependency Network (BDN)-analyse. Dette fordi denne analysen krever at virksomheten finner ut hva

som gir gevinster i virksomheten, og hvordan en implementering av data-drevne makroanalyser eventuelt vil påvirke disse gevinstene. Det er også elementært for å begrunne hva intensjonene bak implementeringen vil være. Dette kan også begrunnes med hva vi kartla i intervjuene. Konklusjonen fra intervjuene var at virksomheter gjerne ikke visste hvordan eller hvorfor de valgte å digitaliseres eller endre interne prosesser, de bare ønsket det fordi de trodde eller hadde hørt at det kunne gi fordeler. Dette kunne være fordeler som prosessoptimalisering eller effektivisering.

Da vi spurte ansatte i twoday Avento «Hva burde vektlegges av bedrifter når det kommer til digitalisering?» var svarene vi gjennomgående fikk «kost-nytte», mulige gevinster og «vil det koste mer enn det smaker?» som respons. I tillegg var responsen at en virksomhet trengte å bli klar over hva digitaliseringen ville innebære i forhold til endringer av prosesser og arbeidsmåter for ansatte og ledelse. Dette ville være element som man internt må gå gjennom før man eventuelt digitaliseres. Derfor konkluderte vi at kartlegging av gevinster i samarbeid med en PESTEL- og SWOT-analyse ville gi et godt grunnlag for evaluering av virksomhetens utgangspunkt i et marked. I tillegg ville analysene til sammen dekke godt utover de åtte dimensjonene man trenger for å følge rammeverket videre. Ved å gjøre det på denne måten kunne vi også konkludere med at rammeverket vårt ville klare målet om å bli en hybrid mellom gevinstrealisering og prosjektledelse. Dette fordi BDN ville kartlegge gevinster, påvirkningen disse har, og hvordan de endres, samt hva konsekvensene av endringene ville bli. PESTEL vil danne et grundig grunnlag for forretningsforståelse – internt og eksternt, og forlengelsen SWOT ville kartlegge viktige styrker og svakheter som til sammen ville være kritiske for at CRISP-DM skal fungere både som en prosess for planlegging, utvikling og implementering, men også som en prosess for endringsledelse i virksomheten. Etter tilbakemelding var mottatt og gjennomgått, måtte vi gå tilbake til steg 3 design og utvikling, for å videreutvikle rammeverket vårt. Innvendinger ble tatt høyde for som beskrevet i evalueringen over. Basert på dette begynte vi skisseringen av vår løsning til rammeverk. Denne vil presenteres i kapittel 5.

### 3.7 Steg 6 Demonstrasjon og evaluering

Her ble det skisserte løsningen igjen demonstrert for oppgavestiller. Dette ble gjort på samme måte som første gang. Rammeverket ble lagt inn i et eget dokument, sammen med en tilhørende forklaring. Dette for å sikre at de som leste gjennom forsto konteksten rundt rammeverket. Basert på dette fikk vi tilbakemelding som videre vil bli presentert i evalueringen.

Basert på mottatte tilbakemeldinger fra demonstrasjonene, foretok vi endringer diskutert i første evalueringsfase. Som vist i figur 3, ble fasen forretningsforståelse i rammeverket, endret til gjennomføring av BDN, PESTEL og SWOT-analyse. Begrunnelsen bak dette, er at disse tre formene for analyse danner et kritisk informasjonsgrunnlag som øker forståelsen av virksomheten og dens omgivelser. En slik forståelse er med på å sikre at videre beslutninger, forankres i virksomhetens reelle behov. Vi mottok positive tilbakemeldinger på denne endringen. Oppgavestiller var enig i at kombinasjonen av disse tre formene for analyse, i høy grad dannet et dekkende og hensiktsmessig informasjonsgrunnlag for å forstå virksomhetens behov. Dette vil igjen gjøre virksomheter bedre rustet til å foreta beslutninger om videre analyse, som vil tilføre dem verdifull innsikt og muliggjøre flere fordeler. Vi diskuterte inkluderingen av andre analysemetoder i denne fasen, men ble enige om at disse tre metodene var dekkende nok for rammeverkets formål. Evalueringen endte i en felles enighet om at vi var

fornøyde med arbeidet i denne fasen, og at dette dermed ville bli en del av det endelige resultatet.

Av første demonstrasjon avdekket vi et ønske fra oppgavestiller om at vi skulle være mer konkrete med tanke på hvilke aktiviteter virksomheter burde gjennomføre i de ulike fasene av rammeverket vårt. I tillegg til dette, var tilbakemeldingen at vi måtte konkretisere hva man skal sitte igjen med etter at de ulike fasene er gjennomført, slik at bruk av rammeverket ikke krever forhåndskunnskap om CRISP-DM. For å konkretisere hva man skal sitte igjen med etter fasene, utarbeidet vi introduksjoner til hver enkelt fase, slik at dette forklares innledningsvis. For å være mer konkrete rundt hvilke aktiviteter virksomheter burde gjennomføre i de ulike fasene, har vi utarbeidet og inkludert nøye beskrivelser av dette i kapittel 5. Tilbakemeldingen vi fikk på arbeidet med disse fasene i rammeverket, var nærmest utelukkende positiv. Det ble ikke påpekt noen spesifikke momenter vi burde forbedre. Basert på tilbakemeldingene fra siste demonstrasjon, samt vår egen forståelse, føler vi at resultatet vi har kommet fram til besvarer problemstillingen vår. Vi anser det dermed ikke som nødvendig å gå tilbake og gjennomføre prosessen en gang til.

## 4 Krav for rammeverk

Etter å ha gjennomført steg 1-4, kom vi fram til et sett med krav til hva som kreves av et rammeverk for å utvikle data-drevne makroanalyser i en virksomhet. I dette kapittelet skal gjøre rede for disse kravene og underbygge de i forhold til intervjudata og eksisterende litteratur. Videre følger en systematisk gjennomgang av kravene til alle de seks fasene i CRISP-CM, sett i lys av Buitelaars åtte dimensjoner. Alle krav har en tilhørende beskrivelse, hvor vi forankrer kravet i litteratur og/eller informasjon hentet fra intervju.

Vi ønsker gjennom et rammeverk å realisere mange av de verdiene som ligger i bruk av data for å drive makroanalyser. Dersom virksomheter oppfyller kravene i rammeverket, vil sannsynligheten for å lykkes med data-drevne makroanalyser øke. Kravene vil altså bidra til endring i virksomheten som er nødvendig for å gjøre den tilpasningsdyktig. Samtidig er det viktig å ta i betraktning at det er stor variasjon blant virksomheter, noe som gjør det vanskelig å utvikle et fasitsvar som vil passe alle. Verdien et rammeverk gir, sammenliknet med for eksempel en metode, vil være at det er mer generelt utformet. Det vil derfor være lettere for andre å relatere til den aktuelle situasjonen, og videre foreta seg endringene som kreves, uten at prosessen blir for kompleks.

Makroanalyser har som formål å skape et detaljert bilde av omgivelsene i en aktuell situasjon. For virksomheter betyr dette at de tilegner seg evnen til å forstå sosiale og økonomiske aspekter på et dypere nivå. Dette innebærer å forstå hvordan aspektene påvirker hverandre i den dagsaktuelle situasjonen, og hvordan de vil utvikle seg i framtiden. Videre vil data-drevne makroanalyser bidra til å forutse potensielle utfall ved en handling, samt hvordan utfallene påvirker omgivelsene rundt de som begår handlingen. Data-drevne makroanalyser vil derfor tilby en ny dimensjon som gir reflekterte og velinformerte beslutninger med en lavere tilknyttet risiko. Basert på historisk data som allerede er innsamlet, vil man utarbeide seg mer presise analyser. Gjennom de tilgjengelige datasettene vil man både kunne forutse fremtidige trender, samt potensielt fremtidige hendelser. Dette kan være alt fra teknologitrender, analyser av utviklingen av økonomi, til større hendelser på verdensbasis.

På sikt kan virksomheter opparbeide seg konkurransefortrinn basert på dette. Data-drevne makroanalyser vil bidra til et bedre beslutningsgrunnlag som kan spare virksomheter for ressurser dersom de riktige avgjørelsene blir tatt. Samtidig vil data-drevne makroanalyser bidra til at man forstår dagens situasjon, hvordan denne utvikler seg, samt hvordan og når man bør handle. I tillegg til å utvikle en forståelse blant virksomheters ytre omgivelser, kan man like gjerne gjennom data-drevne analyser utvikle en forståelse blant virksomheters indre omgivelser. Analyser kan for eksempel synliggjøre ineffektive prosesser, eller lite kostnadseffektive roller, og deretter komme med forslag til en optimalisering eller automatisering av det aktuelle problemet. Med andre ord, vil man i tillegg til å få påpekt problemer, få en rekke alternativer til løsninger – der man også er klar over ettervirkninger og konsekvenser.

Under forskningen til et faglig sterkt og teoretisk rammeverk, vil aktuelle aktører og interessenter bli bevisst på hvordan dagens situasjon i markedet er. Når man lager et rammeverk foretar man seg mye forskning, der man innhenter informasjon fra et bredt

spekter av ulike type kilder. Basert på all informasjonen man får fra de diverse kildene, bør man få en god forståelse for markedet man opererer i. Et rammeverk vil altså kunne blottlegge eventuelle hull i markedet, som for eksempel å påpeke eksisterende løsninger som ikke er optimale. Et resultat av konstruksjonen av et rammeverk, vil være at man kan oppleve ringvirkninger i de aktuelle omgivelsene. For virksomheter vil et tydelig utviklet rammeverk bidra til en økt verdi hos aktører og interessenter. Tydeligheten og omdømme til rammeverket som brukes, vil påvirke hvordan aktører og interessenter ser på virksomheten. Dersom en virksomhet har klart å vedta en god bruk av et respektert rammeverk, vil legitimiteten og aksepten for virksomheten øke. For virksomheten betyr dette at de opparbeider seg en sterkere posisjon i markedet. Ved å ha en sterk posisjon i sitt marked vil man også bli attraktiv å samarbeide med, enten om det er snakk om sponsorer, leverandører, eller aktører som ønsker et samspill.

Med verdigrunnlaget fastsatt begynte vi å se på hvordan utformingen av rammeverket kunne påvirke dette. Vi konkluderte raskt med at det eksisterer mange forskjellige typer rammeverk, og at de alle kunne trekke frem forskjellige styrker eller svakheter fra en virksomhet. Vårt rammeverk har som hovedformål å frigjøre verdien knyttet til data-drevne makroanalyser og vil derfor ta for seg to aspekter ved denne prosessen; en digital endring i en virksomhet, og realisering av verdi. Basert på dette konkluderte vi med at rammeverket vårt ville bli utviklet basert på en hybridløsning av prosjektledelses- og gevinststyringsrammeverk.

I konklusjonen skrevet basert på en deskriptiv analyse, kom Badewi (2022) fram til resultatet om at både prosjektledelse- og gevinststyringsrammeverk har sine fordeler og ulemper. Dersom man sammenligner gevinststyringsrammeverket med prosjektledelsesrammeverket, vil man på et transformasjonsprosjektnivå se at gevinststyringsrammeverket kan framstå svakere og mindre modent enn prosjektledelsesrammeverket. Til tross for dette resultatet, vil gevinststyringsrammeverket ha en sentral rolle for å realisere gevinster, og dermed oppnå transformasjonssuksess (Badewi, 2022). På den andre siden, er prosjektledelsesrammeverk en kritisk faktor for å fullføre vellykkede prosjekter. For å oppsummere, kan man se på et prosjektledelsesrammeverk som essensielt for å tilfredsstille de behovene som de aktuelle prosessene i et prosjekt baseres på. Samtidig vil et gevinststyringsrammeverk være meget viktig for å faktisk få verdi fra et prosjekt, samt at det bidrar til å synliggjøre veien videre for transformasjonsprosessen (Badewi, 2022).

## 4.1 Forretningsforståelse

### Data

#### Krav

For å oppnå forretningsforståelse er man avhengig av aktiviteter som avdekker hvilke data virksomheten innehar, og eventuelt hva som må skaffes

#### Beskrivelse

Virksomheten er avhengig av oversikt over dataene de innehar og at de tar eierskap til dem. Dette vil si at man må gå gjennom eksisterende rutiner for datainnsamling, samt eksisterende data i bedriften for å kartlegge hva man alt har og hva man eventuelt trenger. Når man skal foreta et dataanalyseprosjekt, «*vil man være nødt til å forholde seg til data som allerede eksisterer, at de er tilgjengelige, samt at det foreligger tilstrekkelig grad av detalj slik at de kan brukes*» - (Informant #3). Data kan hentes internt fra mange ulike kilder som for eksempel datavarehus, eller ifølge informant #1 «*interne systemer som ERP og SCM*». Rammeverket må derfor ta for seg aktiviteter knyttet til prosessen rundt kartlegging av internt eksisterende data.

### Metrikker

#### Krav

Den digitale endringen må være må ha klare metrikker for at suksessen kan være målbar.

#### Beskrivelse

IT-bransjen er relativt ung. På grunn av dette er IT-bransjen i stor grad avhengig av å finne sine metrikker selv (Sporsen, 2021). En vurdering av nyttige metrikker i denne fasen, vil derfor være hensiktsmessig å betrakte i skisseringen av rammeverket.

### Ferdigheter

#### Krav

For å oppnå forretningsforståelse er man avhengig av aktiviteter som kartlegger de ansattes ferdigheter og kompetanse.

#### Beskrivelse

Kartlegging av ansattes ferdigheter og kompetanse er ekstremt viktig for å danne en helhetlig forståelse av virksomhetens nå-situasjon. Dette kan gjøres på mange forskjellige måter (Menshikov, 2017). Følgende vil det være hensiktsmessig å inkludere en vurdering av hvordan virksomheter kan gjøre dette på best mulig måte.

### Teknologi

#### Krav

For å oppnå forretningsforståelse er man avhengig av aktiviteter som viser og/eller tester om virksomheten innehar teknologi som understøtter ønsket digital endring

#### Beskrivelse

Dersom virksomheter ønsker å utvikle data-drevne makroanalyser, er de avhengig av teknologi som muliggjør de nødvendige aktivitetene. Teknologien må understøtte «*drift av nødvendig infrastruktur til datavarehus og analyse. Dette kan være i form av datakraft, lagring, lisenser, tilganger, nettverk, integrasjoner*» - (Informant #4). Aktiviteter her vil danne en forretningsforståelse rundt hvilke teknologier virksomheten har tilgjengelig, som igjen danner grunnlaget for en vurdering av hva

som eventuelt kreves av ny teknologi.

## **Ledelse**

### **Krav**

Forretningsforståelsen har sitt grunnlag i ledelsen. Derfor må man ha ledere som er kompetente og tydelige om hvor, hvorfor og hvordan man ønsker å lede bedriften.

### **Beskrivelse**

Dette vil gjenspeiles i de ansattes motivasjon, den generelle kulturen i bedriften og hvordan de kommer til å takle implementeringen.

## **Kultur**

### **Krav**

For å oppnå målet om forretningsforståelse må det gjennomføres en kartlegging av kultur og digital kultur innad i bedriften.

### **Beskrivelse**

Kartlegging av kulturen i bedriften vil være hensiktsmessig for å kartlegge potensiell motstand mot endring innad i virksomheten (Gray, 2017).

## **Strategi**

### **Krav**

Forretningsforståelsen har sitt grunnlag i virksomhetens strategi. Det er derfor viktig at denne er formulert tydelig, målbar og har en generell konsensus i virksomheten.

### **Beskrivelse**

En tydelig formulert strategi vil være med på å forberede de ansatte på at en endring er i vente og kan derfor påvirke kulturen til å være positiv mot endringen.

## **Smidighet**

### **Krav**

Her må det gjennomføres aktiviteter som fastslår hvor tilpasningsdyktig og handledyktig virksomheten er.

### **Beskrivelse**

En virksomhets smidighet fastslås av flere faktorer:

- Ledelse: smidig ledelse fokuserer gjerne på å utvikle andre ledere på flere nivåer for å bedre jobbe mot et felles mål.
- Styring: smidig styring er preget av åpenhet og kontinuitet i hvordan en organisasjons mål settes og oppnås, hvordan risiko håndteres og hvordan ytelse optimaliseres.
- Mennesker: smidige team motiveres av autonomi, mestring og mulighet for å utvikle ferdighetene sine og samarbeidet i teamet.
- Kultur: En smidig kultur kjennetegnes ved et sett kjerneverdier, adferd og praksis som er felles for bedriften. Denne typen kultur fremmer, oppmuntrer og belønner ansattes autonomi.
- Strategi: En smidig strategi utvikles gjennom en iterativ og kontinuerlig prosess. I stedet for å ha én plan på et høyt nivå, lages smidige strategier som et svar på problemer eller hendelser og derfor testes og endres regelmessig.  
(Baker, 2022)



## 4.2 Dataforståelse

### Data

#### Krav

I steget dataforståelse skal rammeverket inneholde aktiviteter som tar for seg hvordan virksomheter kan anskaffe den riktige dataen gjennom tilgjengelige kilder.

#### Beskrivelse

I dataforståelses-fasen skal dataen hentes inn. Utfordringen her er å vite hva som utgjør den riktige dataen, og hvordan virksomheten skal hente den inn. I noen tilfeller vil virksomheter allerede ha den nødvendige dataen, men det vil ofte «*kreve data som ikke allerede har blitt samlet inn før*» - (Informant 3).

Særlig i forhold til bruk av ekstern data, som kan være spesielt relevant i utviklingen av makroanalyser, finnes det utallige kilder å hente inn data fra (Aaser & McElhaney, 2021). Videre følger en vurdering av metode for datainnsamling og eventuelt kjøp av tredjepartsdata. Datainnsamling er avgjørende for å sikre at dataene er samlet på en lovlig og etisk måte og at dataene er fullstendige. Dersom disse to kravene ikke er oppfylt kan det få store konsekvenser for videre bruk av dataene (Cote, 2021). Rammeverket må dermed inneholde aktiviteter som tar for seg aktiviteter knyttet til hvordan virksomheten skal samle inn den riktige dataen for analysens formål.

### Metrikker

#### Krav

Rammeverket må ta for seg hvilke metrikker som kan være nyttige å benytte for å bedre forstå dataen som vurderes.

#### Beskrivelse

Bruker av rammeverket vil få en bedre forståelse av hvilke metrikker som kan være verdifulle i dataforståelsesfasen. I lys av Sporseem (2021), vil det mtp. metrikker innen denne fasen ofte kreve at virksomheter som tar i bruk makroanalyser, utvikler disse selv. «*Sluttproduktene de (virksomheter) er i stand til å produsere gir lite verdi for virksomheten dersom det er mange problemer knyttet til datakvaliteten. Typiske problemer er inkonsekvens, nullverdier og ufullstendig data*» - (Informant 2). En vurdering av hvordan dette kan gjøres gjennom riktig bruk av metrikker, vil derfor være hensiktsmessig å ta med.

### Ferdigheter

#### Krav

I steget dataforståelse skal rammeverket inneholde en oversikt over hva som kreves av kunnskap og ferdigheter innen data, databehandling og forretningsforståelse, for at virksomheten effektivt skal kunne hente inn den riktige dataen.

#### Beskrivelse

Det er svært viktig å etablere en detaljert forståelse av hvilke kunnskaper og ferdigheter som er relevant for virksomhetens ansatte, for å sikre at riktige personer blir satt til, evt. ansatt og opplært, og følgende at riktig data blir hentet inn og tatt videre i prosessen (Radovilsky et al., 2018). Det vil derfor være nyttig å etablere hva som utgjør relevant kunnskap og ferdigheter i denne fasen av et makroanalyseprosjekt, og hvordan virksomheten kan sikre det.

## **Teknologi**

### **Krav**

I steget dataforståelse skal rammeverket ta for seg hvilke komponenter en virksomhets dataarkitektur må bestå av, for å muliggjøre innhenting, lagring og tilgjengeliggjøring av nødvendig data.

### **Beskrivelse**

Avhengig av virksomhetens strategi og mål bak makroanalyser, vil ulike teknologi møte deres behov. I denne fasen, må dataen hentes inn og bli gjort tilgjengelig for ansatte som skal utvikle analysen. Dette vil kreve ulike teknologiske løsninger. Ifølge informant 2, vil «*løsningene som er tilgjengelig, drøftes og diskuteres med kunden. Avhengig av kompleksitet i hva kunden etterspør, vil for eksempel fullstack-løsninger eller direkte rapporter/dashboards i PowerBI diskuteres*». Ulike virksomheter har dermed ulike behov for teknologi i denne fasen. Rammeverket skal gi bruker en oversikt over hva som kreves rent teknologisk for å utvikle data-drevne makroanalyser.

## **Ledelse**

### **Krav**

Rammeverket skal inneholde en vurdering av hva som kreves av ledelse i dataforståelsesfasen av et makroanalyseprosjekt.

### **Beskrivelse**

God ledelse er essensielt i et hvert prosjekt, for at målene med prosjektet skal oppfylles. Følgende vil det være viktig å avgjøre hva som utgjør god ledelse i denne fasen av et makroanalyseprosjekt, for at virksomheten skal anskaffe best mulig datagrunnlag før påbegynnelse av neste fase.

## **Kultur**

### **Krav**

Rammeverket skal inneholde en oversikt over virksomhetens interne kulturelle forhold som kan påvirke graden av suksess i fasen dataforståelse.

### **Beskrivelse**

Organisasjonskultur kan ha stor betydning i forhold til hvor verdiskapende interne prosesser er for virksomheten, også når det kommer til digital-/datakultur og dataanalyseprosesser (Martinez-Caro et al., 2020). Det vil derfor være viktig å ta for seg interne kulturelle forhold som kan påvirke denne i henhold til fasen dataforståelse.

## **Strategi**

### **Krav**

Rammeverket skal ta for seg hva som kreves av virksomhetens strategi, for å muliggjøre innsamling/ibruktagelse av hensiktsmessig data.

### **Beskrivelse**

Avhengig av virksomhetens mål og strategi, kan ulike data benyttes til å besvare forretnings spørsmålene. Hvilke data som vil være hensiktsmessig, «*kommer an på hva virksomheten ønsker å benytte dataen til*» - (Informant #1). Rammeverket må dermed ta for seg hva som kreves av virksomhetens strategi, for å muliggjøre dette på tvers av virksomheten.

## Smidighet

### Krav

Rammeverket må inneholde løsninger for hvordan dataarkitekturen kan holdes oppdatert og relevant for å møte mulige eksterne og interne endringer.

### Beskrivelse

Den største utfordringen knyttet til smidighet er ifølge informant #4 «å holde koblinger oppdatert og fange opp når f.eks. datakilder endres eller slutter å svare. Løpende vedlikehold og overvåkning er dermed veldig viktig siden slike systemer gjerne har mange komponenter og er sårbare for endringer». Rammeverket må derfor ta seg hvordan virksomheter kan forholde seg smidige i denne fasen av et makroanalyseprosjekt.

## 4.3 Dataforberedelse

## Data

### Krav

I steget dataforberedelse må rammeverket inneholde aktiviteter som tar for seg hvordan man forbereder data på en hensiktsmessig måte for å foreta makroanalyser.

### Beskrivelse

En av de store utfordringene innen dataanalyse er nettopp det å rense dataen til valgt formål (Chu et al., 2016). Data må forberedes på ulike måter avhengig av bruksformålet med dataen, innsikt man ønsker fra den, samt karakteristikker ved dataen i seg selv. Rammeverket må inneholde en forklaring av hvordan dette kan gjøres mest hensiktsmessig, da det er kritisk for bruk av dataen man tar for seg i de følgende fasene.

## Metrikker

### Krav

Rammeverket må ta for seg ulike metrikker som kan være nyttige å benytte i fasen dataforberedelse.

### Beskrivelse

Dataforberedelse er ofte den mest tidskrevende fasen av data science prosjekter (IBM, 2021d). Det ligger derfor mye verdi i det å optimalisere prosesser i denne fasen. Gode metrikker kan være et viktig hjelpemiddel å benytte her. Det vil også være hensiktsmessig å benytte metrikker for å vurdere andre resultater av dataforberedelsesprosessen knyttet til datakvalitet, samt overvåke endrende data (Stedman, 2020). En diskusjon rundt utvikling av hensiktsmessige metrikker, må derfor inkluderes i rammeverket.

## Ferdigheter

### Krav

I steget dataforberedelse må rammeverket inneholde en oversikt over ferdigheter og kunnskap som kreves for å forberede dataen til riktig formål.

### Beskrivelse

Det vil være viktig å forstå hva som kreves av kunnskap og ferdigheter også i dette steget. De ansatte som får oppgaver innen dataforberedelse, må være i stand til å utføre disse på en hensiktsmessig måte. I tillegg blir dataanalyse i økende grad

vesentlige viktigere i virksomheter, og følgende også et behov for at ansatte forstår, deler kunnskap rundt, og har meningsfulle diskusjoner/samtaler om data (Panetta, 2021). For å sikre at riktige ansatte blir satt til riktige oppgaver, og at ansatte generelt innehar nødvendig kunnskap rundt data, vil rammeverket inneholde en vurdering av relevant kunnskap og ferdigheter i denne fasen.

## **Teknologi**

### **Krav**

I steget dataforberedelse skal rammeverket ta for seg hvilke komponenter en virksomhets dataarkitektur må bestå av for å muliggjøre håndtering og forberedelse av data.

### **Beskrivelse**

Lik foregående fase vil riktig teknologi også her avhenge av varierende faktorer knyttet til virksomheten, og formålet deres med data-drevne makroanalyser. Rammeverket skal ta for seg hva dette innebærer, og hvilke teknologiske komponenter som kreves for å forberede dataen til valgt formål.

## **Ledelse**

### **Krav**

Rammeverket må ta for seg hva som kreves av ledelse i dataforberedelsesfasen.

### **Beskrivelse**

God ledelse er essensielt i et hvert prosjekt. Følgende vil det være viktig å avgjøre hva som utgjør god ledelse i denne fasen av et makroanalyseprosjekt.

## **Kultur**

### **Krav**

Rammeverket skal inneholde en oversikt over virksomheters interne kulturelle forhold som kan påvirke graden av suksess i fasen dataforberedelse.

### **Beskrivelse**

En virksomhets kultur kan ha stor betydning på deres grad av suksess i henhold til Martinez-Caro et al. (2020). Det vil derfor være viktig å ta for seg interne kulturelle forhold som kan påvirke denne i henhold til fasen dataforberedelse.

## **Strategi**

### **Krav**

Rammeverket skal ta for seg hva som kreves av virksomheten for å muliggjøre en dataforberedelse forenelig med virksomhetens strategi

### **Beskrivelse**

Hvilke data som vil være hensiktsmessig å benytte, «*kommer an på hva virksomheten ønsker å benytte dataen til*» - (Informant #1). Dette har også utslag i hvordan dataen forberedes til analyseformål. Virksomhetens strategi står sentralt i å avgjøre hva dataen skal benyttes til. En diskusjon av hva dette innebærer i henhold til strategi, vil derfor inkluderes i rammeverket.

## **Smidighet**

### **Krav**

Rammeverket må ta for seg hva som kreves av smidighet i dataforberedelsesfasen.

### **Beskrivelse**

For å effektivt utvikle makroanalyser, krever det at virksomheten klarer å tilpasse analyseprosessen med tanke på endringer som kan oppstå underveis. Dette kan i henhold til Informant #4 være endringer i datakilder, men også endringer forårsaket av f.eks. endret dataforståelse underveis i prosjektet. I denne fasen, må virksomheten også være i stand til å ta høyde for interne og eksterne endringer som kan oppstå under utviklingen av makroanalyser, og følgende må rammeverket ta for seg hvordan dette kan gjøres.

## 4.4 Modellering

### **Data**

#### **Krav**

Rammeverket må inkludere en beskrivelse av aktiviteter for hvordan virksomheter best kan velge modelleringsteknikk.

#### **Beskrivelse**

Dette må gjøres for at modellen skal ha det beste utgangspunktet og bli så nøyaktig som mulig i forhold til bruksmålet. I henhold til overnevnte krav i tidligere faser, er formålet med dataanalysen essensielt å betrakte også her.

### **Metrikker**

#### **Krav**

Rammeverket må ta for seg hvilke metrikker som er hensiktsmessig for å måle for å finne ut om modelleringen og modellen bidrar positivt til virksomheten.

#### **Beskrivelse**

Metrikker kan være et viktig verktøy å benytte i utviklingen og spesielt vurderingen av ulike modeller i denne fasen. I lys av Sporsem (2021), vil det ofte kreve at virksomheter ofte utvikler disse selv også i denne fasen. Rammeverket må derfor inneholde en diskusjon rundt dette.

### **Ferdigheter**

#### **Krav**

Rammeverket må ta høyde for om modelleringen skal skje ved bruk av mennesker og hvilke ferdigheter de må inneha.

#### **Beskrivelse**

Modeller kan utvikles gjennom bruk av mange ulike teknikker, men ansvaret for å velge de riktige teknikkene vil i utgangspunktet ligge hos de involverte ansatte (Bihani & Patil, 2014). Det er dermed viktig at de ansatte som involveres innehar de riktige ferdighetene, for å gjennomføre fasen på en måte som fører til verdifull innsikt. Av denne grunn vil vi inkludere en diskusjon rundt hva riktige ferdigheter utgjør. Dette for å sikre at de riktige ansatte involveres, eventuelt for å kartlegge hva virksomheten må bruke av ressurser for å ansette eller lære opp ansatte til å gjennomføre modelleringen.

|                  |   |
|------------------|---|
| <b>Teknologi</b> | <p><b>Krav</b><br/>Rammeverket må ta høyde for hvilke teknologier ulike modelleringsteknikker behøver.</p> <p><b>Beskrivelse</b><br/>Ulike modelleringsteknikker kan kreve ulike former for teknologi (Bihani &amp; Patil, 2014). Rammeverket må inneholde en forklaring av dette for å kartlegge hva virksomheten eventuelt må bruke av ressurser for å skaffe eller oppgradere teknologi, slik at de er i stand til å gjennomføre ønsket modellering.</p>   |
| <b>Ledelse</b>   | <p><b>Krav</b><br/>Rammeverket må inneholde aktiviteter som kan bistå ledelsen i arbeidet med å overse og forsikre at modelleringen skjer i samsvar med virksomhetens strategi.</p> <p><b>Beskrivelse</b><br/>For å sikre forretningsverdi fra analyseprosjektet er det viktig at modelleringen skjer i samsvar med virksomhetens overordnede strategi og mål. God ledelse vil være viktig her, og følgende må rammeverket inneholde en diskusjon rundt hva dette innebærer i denne fasen av et makroanalyseprosjekt.</p> |
| <b>Kultur</b>    | <p><b>Krav</b><br/>Rammeverket må ta høyde for type kultur i virksomheten, og hva som kreves av den i denne fasen.</p> <p><b>Beskrivelse</b><br/>En virksomhets kultur kan ha stor betydning på deres grad av suksess i henhold til Martinez-Caro et al. (2020). Det vil derfor være viktig å ta for seg interne kulturelle forhold som kan påvirke denne i henhold til fasen modellering.</p>  |
| <b>Strategi</b>  | <p><b>Krav</b><br/>Rammeverket må bidra til at modelleringen skjer i henhold til strategien virksomheten har satt.</p> <p><b>Beskrivelse</b><br/>Det er viktig at analysen bidrar til verdifull innsikt forenelig med virksomhetens strategi. Rammeverket må dermed ta for seg hva som kreves for at dette skal være gjennomførbart.</p>  |
| <b>Smidighet</b> | <p><b>Krav</b><br/>Rammeverket må ta for seg hva som kreves av smidighet i modelleringsfasen</p> <p><b>Beskrivelse</b><br/>I henhold til informant #4 sin uttalelse under beskrivelse av krav til smidighet i dataforståelsesfasen, kan det oppstå mange utfordringer knyttet til ulike eksterne og interne endringer også her. Rammeverket må inkludere en forklaring av aktiviteter knyttet til å møte slike utfordringer i denne fasen.</p>  |

## 4.5 Evaluering

### Data

#### Krav

Rammeverket må ta for seg hvordan dataen skal evalueres for å sikre god datakvalitet.

#### Beskrivelse

En god datakvalitet vil gi en nøye oversikt over hva man har foretatt seg tidligere, og vil derfor bidra til å forstå den dagsaktuelle situasjonen bedre. Det legger et grunnlag for aktuelle avgjørelser og videre analyser (Chapman, 1999).

### Metrikker

#### Krav

Rammeverket må ta for seg hva som kreves av metrikker for å vurdere modellene, og avgjøre og om disse er tilfredsstillende. Dersom de møter virksomhetenes standarder, kan de implementeres.

#### Beskrivelse

Ved å foreta seg en vurdering av modellene, vil man få en pekepinn på alt som er gjort, om dette er godt nok, og deretter legge et grunnlag for hva man bør gjøre i framtiden. Ved å gjøre de godkjente modellene om til data, vil man anskaffe seg et større datagrunnlag for videre analyser o.l.

### Ferdigheter

#### Krav

Rammeverket må ta for seg hva som kreves av ferdigheter for å gjennomføre forsvarlige evalueringer.

#### Beskrivelse

Ferdigheter og relevant erfaring er en essensiell ferdighet når man foretar seg evalueringer. Basert på tidligere beslutninger kan man opparbeide seg kunnskap om hva som tradisjonelt har fungert godt, og ikke fullt så godt (Chapman, 1999). Denne kompetansen bidrar til et bedre beslutningsgrunnlag, samt en mer reflektert evaluering for virksomheten.

### Teknologi

#### Krav

Rammeverket må ta for seg hva som kreves av teknologi for å evaluere resultatene.

#### Beskrivelse

For å kunne gjennomføre en grundig evaluering bør virksomheten ha tilstrekkelig med aktuelle verktøy, som analyseverktøy, visualiseringsverktøy, verktøy for modelltesting, o.l. Slike verktøy optimaliserer, automatiserer og effektiviserer aktuelle prosesser innad i virksomheten. Som et resultat av optimaliseringen, automatiseringen, og effektiviseringen, kan virksomheten frigjøre eller omplassere midler og ressurser. Dette kan lede virksomheten til blant annet økonomiske gevinster, økt arbeidskapasitet, eller effektivisering blant prosesser innad i virksomheten, da de er kapable til å effektivt omplassere ressurser basert på foreliggende behov.

### Ledelse

#### Krav

Rammeverket må ta for seg hva som kreves av ledelse, under evalueringsfasen av prosjektet.

### **Beskrivelse**

Ledelsen bør tenke langsiktig, og tidlig legge en plan for videre steg virksomheten skal ta. Ved å tenke langsiktig, vil man tilføye virksomheten en stabilitet. Man vil vite hvordan man skal handle, og unngå kaotiske situasjoner (Chapman, 1999). Med andre ord, vil man oppnå en forutsigbar situasjon for virksomheten. En liste for aktuelle handlinger virksomheten kan ta i framtiden bør vurderes.

## **Kultur**

### **Krav**

Rammeverket må ta for seg hva som kreves av kultur for å sikre en vellykket evaluering av de produserte resultatene.

### **Beskrivelse**

Mange kulturelle faktorer kan påvirke evalueringsfasen. Kulturen bør bl.a. vektlegge håndfaste rutiner for evaluering. Ved å vektlegge håndfaste rutiner for evaluering, vil man oppleve tydelige gevinster. Gevinster som man muligens kan oppleve er nøyaktige resultater, bedre kvalitet ved tjenester og produkter, og et legitimt og seriøst omdømme blant interessenter og aktører. Ved å vise at man vektlegger rutiner for evaluering tungt, vil man vise seg som en seriøs aktør på markedet, da evalueringer er et svært viktig steg for å lykkes. Samtidig vil håndfaste rutiner også bidra til at man til enhver tid vet hva situasjonen krever, og har dermed steg å forholde seg til. Dette leder for øvrig til at produktet eller tjenesten som produseres er nøye gjennomført og gjennomtenkt, som bidrar til en økt kvalitet.

## **Strategi**

### **Krav**

Rammeverket må ta for seg hva som kreves av strateg

### **Beskrivelse**

Slik som nevnt under ledelse, bør man legge en plan for hva som skal gjøres videre etter en evaluering. Dette avhenger åpenbart av resultatene fra evalueringen. Eksempel kan være å revidere enkelte faser som ikke er optimale, eller ta steget videre til neste fase (Chapman, 1999).

## **Smidighet**

### **Krav**

Rammeverket må ta for seg hva som kreves av smidighet under evalueringsfasen

### **Beskrivelse**

Basert på den foretatte evalueringen, bør man gjøre opp en vurdering på hvorvidt resultatene møter virksomhetens behov. Hvis ikke bør man gjøre eventuelle justeringer og endringer slik at den faktisk møter virksomhetens reelle behov. Dersom modellens formål ikke møter virksomhetens behov, vil den heller ikke ha stor verdi for virksomheten. Derfor er det viktig å foreta seg aktuelle endringer og justeringer, slik at modellen faktisk gir verdi. Dette må derfor diskuteres i rammeverket.



## 4.6 Implementering

### Data

#### Krav

Rammeverket må ta for seg hva som skal gjøres med dataen som blir oppsamlet under implementeringen.

#### Beskrivelse

I et slikt rammeverk vil data være verdifullt, fordi det har et stort bruksområde. Derfor vil det være aktuelt å utvinne mest mulig forsvarlig data i enhver mulig situasjon (Informant #3). Den utvinnete dataen kan brukes i videre analyser, som for eksempel å predikere rammeverkets fremtidige utvikling.

### Metrikker

#### Krav

Rammeverket må ta for seg hva som kreves av testing og målinger for å unngå potensielle svakheter, for å overvåke at modellen fungerer som forventet, og at den tilfredsstillir virksomhetens behov.

#### Beskrivelse

For å sikre at implementeringen faktisk møter virksomhetens behov, og at den leverer optimale resultater, bør den tidlig testes etter implementeringen. Dermed vil man få oversikt over hva som fungerer, og eventuelt hva som ikke fungerer.

### Ferdigheter

#### Krav

Rammeverket må ta for seg hva som kreves av ferdigheter for å sikre at implementeringen blir vellykket.

#### Beskrivelse

For å kunne gjennomføre tester av implementeringen, er man nødt til å ha et grunnlag til å foreta seg testene. I dette tilfellet vil grunnlaget dannes gjennom ferdigheter, hovedsakelig kompetanse og erfaring. De nevnte ferdighetene er essensielle for å kunne foreta seg vellykkede implementeringer, derav å få et best mulig resultat, samt utbytte fra dem. Samtidig bidrar erfaring og kompetanse til sikkerhet og effektivisering av implementeringen, da man til enhver tid vet hvordan man skal handle, og hvilke valg som bør gjøres (Chapman, 1999).

### Teknologi

#### Krav

Rammeverket må ta for seg hva som kreves av teknologi for å støtte opp implementering av løsningen.

#### Beskrivelse

Teknologi kan anvendes som verktøy som gjør implementeringen enklere, samt optimaliserer prosessen. Slike verktøy være selve implementeringen og testingen av modellen, overvåkning og vedlikehold av de aktuelle prosessene, maskinlæreverktøy, automatiseringsverktøy, eller verktøy som oppbevarer og behandler innsamlet data

### Ledelse

#### Krav

Rammeverket må ta for seg hva som kreves av ledelsen for å kontinuerlig overvåke implementeringen, og deretter planlegge videre arbeid dersom det oppstår bestemte

behov.

### **Beskrivelse**

Overvåkning gir en forståelse av hvordan rammeverket fungerer akkurat nå, samtidig som at man kan se utviklingen den har hatt. Basert på utviklingen som allerede har tatt sted, vil man kunne gjøre analyser av hvordan rammeverket vil utvikle seg i framtiden, og dermed tilegne seg en forutsigbar situasjon (Chapman, 1999). Ved en kontinuerlig overvåkning vil man også ha forutsetning til å raskt gjøre beslutninger dersom det skulle oppstå en avvikende situasjon. Dermed sikrer man seg at man rammeverket alltid yter best.

## **Kultur**

### **Krav**

Rammeverket må ta for seg hva som kreves av opplæring av ansatte i denne fasen.

### **Beskrivelse**

For at ansatte skal kunne ta i bruk rammeverket, er de nødt til å ha en forståelse av rammeverkets formål, hvordan det fungerer, og hvordan de skal ta det i bruk. En åpen kultur der ansatte får en grundig opplæring i rammeverket, der de kan stille hverandre spørsmål, vil ha positive effekter på virksomheten (Al-Mashari & Zairi, 1999). Positive effekter som virksomheten kan oppleve varierer fra trivsel på arbeidsplassen blant de ansatte, til effektivisering blant resultater, da de ansatte vet hvordan de skal handle.

## **Strategi**

### **Krav**

Rammeverket må ta for seg hva som kreves av strategi for effektiv implementering av innsikt/modell.

### **Beskrivelse**

Dersom man har en tydelig og klar strategi for hvordan man skal handle for å unngå avvikende situasjoner, vil man handle med lav risiko. Dette gjøres gjennom å kontinuerlig skåne virksomheten for potensielle farer eller fallgruver (Mabert et al., 2003). Hvis man mot formodning skulle oppleve avvikende situasjoner, vil man gjennom definerte tiltak ha kompetanse og erfaring til å effektivt manøvrere seg vekk fra situasjonen.

## **Smidighet**

### **Krav**

Rammeverket må ta for seg hva som kreves av smidighet dersom uforutsette situasjoner som kan kreve endringer skal oppstå

### **Beskrivelse**

Dersom uforutsette situasjoner skulle oppstå, vil man være nødt til å foreta seg justeringer og endringer. Disse vil ha som formål å tilpasse virksomheten på best mulig måte, slik at man i best grad møter virksomhetens strategi (Chapman, 1999).

## 5 Skisse av rammeverk

Vår skisse av rammeverket vårt er bygd opp CRISP-DMs prosessmodell for gjennomføring av datavitenskaplige prosjekter. Denne tar for seg prosjektets livsløp fra start til slutt, samtidig som den gir virksomheter som ønsker å ta i bruk rammeverket en oversikt over de ulike fasene de må igjennom. De åtte dimensjonene ble inkludert i rammeverket da de tar for seg hovedelementene en virksomhet består av, og er dekkende for både tekniske og organisatoriske krav. Dette i tillegg til at de fungerer utfyllende sammen, siden de åtte dimensjonene er dekt i CRISP-DM modellen. BDN, PESTEL og SWOT ble tatt inn i rammeverket da man kan se at disse til sammen vil gi et godt grunnlag for forretningsforståelse, og greier å skaffe grundig informasjon om alle de åtte dimensjonene. I tillegg tar analysene for seg både mikro- og makroomgivelser som vil være med på å dekke behovet for tekniske og organisatoriske faktorer.

Etter at forretningsforståelse er etablert, vil de resterende fasene i rammeverket, fra dataforståelse til implementering, følge lik struktur. I disse fasene tar vi for oss ulike momenter som må vurderes og/eller aktiviteter som må gjennomføres i henhold til hver av de åtte virksomhetsdimensjonene. Dette for at virksomheter best mulig skal være i stand til å utvikle data-drevne makroanalyser. Totalt sett er rammeverket vårt bygd opp av disse ulike metodene, dimensjonene og analysene fordi de utfyller hverandre. Til sammen skaper de nok informasjon for en virksomhet til å skulle kartlegge nåværende tekniske og organisatoriske faktorer, samtidig som de kan stake ut en ønsket retning for virksomhetens ideelle, fremtidige situasjon.

**Figur 3 Rammeverk for tekniske og organisatoriske krav for utvikling av data-drevne makroanalyser**

| Forretningsforståelse | Dataforståelse     | Dataforberedelse  | Modellering   | Evaluering   | Implementering   |   |
|-----------------------|--------------------|---|---|--|--|---|
| BDN                   | <b>Data</b>        | Tilgjengelig data må velges ut basert på forretningsbehov         | Dårlig datakvalitet må oppdages og fikses   | Riktig modelleringsteknikk(er) på velges ut, basert på formål og dataene i seg selv      | Datakvalitet må evalueres  | Ha tillit til data, og etablere plan for videre steg                      |
|                       | <b>Metrikker</b>   | Metrikker må utvikles for å forstå dataen bedre                   | Det må utvikles metrikker for å styre datakvalitet og prosessen rundt dataforberedelse  | Metrikker for å måle effektivitet i fasen må etableres                                   | Evaluere modellresultatene ved hjelp av metrikker                                  | Måle effekten av implementeringen   |
|                       | <b>Ferdigheter</b> | Det må sørges for riktig, relevant kompetanse                     | Involverte ansatte må inneha relevante tekniske ferdigheter og organisatorisk forståelse  | Ansatte må ha kompetanse innenfor data-drevne makroanalyser og verktøy                   | Relevante ferdigheter for å evaluere prosessen må være på plass                    | Ferdigheter innenfor relevante områder                                    |
| PESTEL                | <b>Teknologi</b>   | ETL-verktøy eller tilsvarende, må implementeres/være implementert | ETL-verktøy (eller tilsvarende) og evt. supplementære verktøy må implementeres/være implementert                                | Modelleringsverktøy må implementeres/være implementert                                   | Virksomhetens infrastruktur bør evalueres, samt den produserte modellen            | Teknologi for å overvåke modell / visualisere data                        |
|                       | <b>Ledelse</b>     | Roller med relevante ansvarsområder må etableres                  | Spesialiserte fagpersoner må være involverte, og en opplæring i generell datakompetanse bør foretas i virksomheten om nødvendig | Ledelsen har ansvar for å overse at modelleringen skjer i samsvar med strategi og formål | Ledelsen må sette mål for fasen, og sørge for gjennomførelse                       | Overvåkning og vedlikehold, samt endringsledelse                          |
| SWOT                  | <b>Kultur</b>      | Kulturen bør fremme innovasjon                                    | Datakvalitetskultur må opparbeides  | Kulturen bør være kritisk og reflektert  | En kritisk, læringsvillig og datakvalitetsorientert kultur bør oppbygges           | Tilstrekkelig opplæring og en læringsvillig kultur som ser verdien i data |
|                       | <b>Strategi</b>    | Strategien må være tydelig definert og kommunisert                | Strategien må være tydelig definert og kommunisert  | Strategien må være tydelig definert og kommunisert                                       | Neste skritt må planlegges. Gå videre til implementering eller tilbake i prosessen | Ekspertise i datavitenskap, og data-drevne makroanalyser                  |
|                       | <b>Smidighet</b>   | Data/datakilder må forvaltes og overvåkes                         | Smidige prosesser og IT-/dataarkitektur   | Nøyaktige modellbeskrivelser av resultatene, mulighet til å gå tilbake i prosessen       | Eventuelle endringer må foretas  | Ha en etablert plan for utfasing av modell                                |

## 5.1 Forretningsforståelse

Den første fasen i rammeverket er forretningsforståelse. Her vil målet være å etablere en god forståelse over hva nå-situasjonen er for virksomheten. Dette både med tanke på hvordan de plasserer seg i et marked, hvem som er konkurrenter, hvilke makroomgivelser de har å forholde seg til og hvordan virksomheten genererer verdi. Naturlig vil det da være å gjennomføre flere analyser som tar for seg flere fagfelt. Dette burde gjøres i et samarbeid mellom ledelse og ansatte som innehar kunnskap innenfor de respektive fagfeltene, samt kan virksomhetene diskutere om de ønsker å inkludere interessenter i denne fasen.

### 5.1.1 Benefit Dependency Network – analyse

Utarbeidelsen av en realiseringsplan begynner med å svar syv spørsmål om gevinstrealisering (Peppard & Ward, 2007). Disse spørsmålene har fokus på organisasjon og business og ikke på IT:

1. Hvorfor må virksomheten forbedre seg?
2. Hvilke forbedringer er nødvendige eller mulige?
  - a. Nøkkelinteressenter må godta disse forbedringene, da disse blir grunnlag for investering
3. Hvilke gevinster vil bli realisert av hver enkel interessent dersom investeringsmålene er nådd?
  - a. Hvordan vil gevinstene bli målt?
4. Hvem er ansvarlig for hver gevinst og vil være ansvarlig for realiseringen?
  - a. Gevinst-eieren vil være ansvarlig for verdien tildelt gevinsten.
5. Hvilke endringer er nødvendige for å oppnå hver enkel gevinst?
  - a. Nøkkelen til å realisere gevinster er å identifisere sammenhenger mellom gevinster og nødvendig endring/forbedring.
6. Hvem vil være ansvarlig for å sikre at endringene er gjennomført vellykket?
7. Hvordan og når kan de identifiserte endringene være gjort?
  - a. For å svare på dette spørsmålet må virksomheten vurdere hver interessentgruppes evne og kapasitet til å gjennomføre de identifiserte endringene.

Det neste vil være å identifisere hvilken type investering dette vil være for virksomheten. Hovedsakelig her skiller vi mellom en problemsentrert eller en innovasjonsbasert tilnærming.

#### 5.1.1.1 Problemsentrert tilnærming

I problembaserte intervensjoner vil en organisasjon typisk investere i IT for å forbedre ytelsen for å (Peppard & Ward, 2007):

- Overvinne en eksisterende ulempe mot konkurrenter.
- Forhindre at ytelsen blir dårligere i fremtiden til et nivå som vil sette virksomheten i en konkurranseulempe.
- Oppnå ønskede forretningsmål.
- Fjerne begrensninger som hindrer muligheter fra å bli tatt.

### 5.1.1.2 Innovasjonsbasert tilnærming

I en innovasjonsbasert tilnærming vil en virksomhet investere i IT for å utnytte en forretningsmulighet, for å skape potensielle konkurransemuligheter eller for å bygge nye organisasjoner ved å (Peppard & Ward, 2007):

- Gjøre noe nytt gjennom bruken av IT.
- Gjøre noe på en nyt måte ved hjelp av IT.
- Bruke ny IT til å gjøre noe virksomheten tidligere ikke har kunnet gjøre.

I alle disse situasjonene vil innovasjonen avhenge av kombinasjonen av teknologi, virksomhetens tekniske ekspertise og virksomhetens evne til å utnytte de nye kapasitetene effektivt.

### 5.1.2 PESTEL-analyse

En PESTEL-analyse er ment til å gi en oversikt og bedre forståelse av langsiktige trender. Dette er en aktivitet som kan støtte en rekke planleggingssituasjoner for virksomheter. Hvilke aktiviteter er listet opp i tabellen under:

**Tabell 3 Formål ved en PESTEL-analyse**

|  |  |
|--|--|
| <b>Strategisk forretningsplanlegging</b>           | PESTEL-analyse gir en kontekstuell informasjon om forretningsretningen, merkevareposisjonering, vekstmål og risikoer (eksempelvis en ny pandemi) for produktiviteten. Dette vil bidra til å bestemme levetiden til et eksisterende produkt eller tjenester og eventuelt gi en indikator på når ny produktutvikling burde skje. |
| <b>Arbeidsstyrkeplanlegging</b>                    | Analysen kan bidra til å identifisere forstyrrende endringer i forretningsmodeller som kan påvirke det fremtidige arbeidslandskapet. Den kan identifisere kompetansegap/mangler, nye jobbroller, jobbreduksjoner eller forskyvinger.   |
| <b>Markedsplanlegging</b>                          | Klimaelementet her kan være essensielt i forhold til markedsføringsprosessen. Det kan bidra til å prioritere forretningsaktiviteter for å oppnå spesifikke markedsføringsmål innenfor en fastsatt tidsramme.   |
| <b>Produktutvikling</b>                            | Ved å overvåke makroomgivelser kan man finne ut om det er lurt å gå inn i eller forlate en del av et marked, eller finne ut om et produkt eller tjeneste fortsatt oppfyller et behov i markedet eller om man burde utvikle eller lansere et nytt produkt.  |
| <b>Organisasjonsendring</b>                        | En PESTEL-analyse hjelper til med å forstå konteksten for en endring, og er mest effektiv når den brukes med en SWOT-analyse for å forstå muligheter og trusler endringer  |
| <b>Menneskestrategier, rapporter og prosjekter</b> | Hovedformålet ved PESTEL-analysen er at den kan brukes som et rammeverk for å se hvilke makroomgivelser som må utforskes videre og prøve å predikere hva som kan skje i fremtiden. Dette kan sikre at viktige faktorer ikke blir oversett når man tilpasser strategier.  |

### 5.1.3 Hvordan gjennomføre en PESTEL-analyse

Gjennomføring av en PESTEL-analyse krever forståelse og forskning for å samle inn informasjonen som er nødvendig. Analysene av de eksterne miljøet må gjøres manuelt, slik at når virksomheten har samlet inn nødvendig informasjon kan disse analyseres for å identifisere virksomhetens neste handlinger (Reding, 2021).

Selve gjennomføringen av PESTEL-analysen burde begynne med å identifisere omfanget av forskningen. Denne identifiseringen burde ta for seg nåværende og fremtidige plausible scenarier. Den burde også omfatte virksomhetsbransjen, markedet og de områdene av verden der virksomheten opererer i.

Videre må det bestemmes hvordan informasjonen skal samles inn. Samle en gruppe ansatte som kan bidra til å samle inn data og synliggjøre bevis og perspektiver. Dersom man har flere personer til å bidra til informasjonsinnhenting, fører dette ofte til rikere data blir fanget opp (Reding, 2021). Innledende ideer kan genereres gjennom brainstorming eller lage et PESTEL-tavle.

Videre må gruppen identifisere passende informasjonskilder, og begynne informasjonsinnhenting, for hver PESTEL-kategori. Virksomhetene kan ha områder som er mer viktige eller aktuelle enn andre, men det vil likevel være hensiktsmessig å utforske alle områdene for å få et større innsyn i virksomhetens makroomgivelser. Under følger en oversikt over spørsmål som kan være viktig å stille seg til de ulike områdene PESTEL-analysen tar for seg (Reding, 2021):

1. Politisk
  - a. Vurder nasjonal og internasjonal politikk (kun hvis dette er aktuelt for din virksomhet)
  - b. Er det politiske miljøet for din virksomhet stabilt?
  - c. Finnes det finansiering eller tilskudd tilgjengelig for virksomheten?
2. Økonomisk
  - a. Er inntekten til virksomhetens kunder i endring? I så fall hvordan?
  - b. Er virksomhetens produksjonskostnader synergisk med produktterspørsel og fortjeneste?
  - c. Hvilke faktorer kan forbedre eller redusere forretningsutsiktene til virksomheten.
3. Sosiologisk
  - a. Hvilken påvirkningskraft har kundeatferd og holdninger på kjøpsvaner?
  - b. Er holdningene til virksomhetens interessenter i endring?
  - c. Hvordan er aldersfordelingen i målgruppen?
4. Teknologisk
  - a. Finnes det ny programvare som kan hjelpe virksomheten?
  - b. Vil den nye teknologien eventuelt true kompetansen til virksomhetens ansatte?
  - c. Brukes det automatisering i markedet virksomheten opererer i? I så fall hvordan brukes denne?
5. Bærekraft
  - a. Hvordan endrer det fysiske miljøet seg?
  - b. Hvordan reagerer virksomhetens kunder på eventuelle endringer?
  - c. Kan virksomheten ta i bruk andre eller bedre materialer i produksjon av varer og tjenester?
6. Juridisk

- a. Hvilke lover og regler må virksomheten forholde seg til?
- b. Vil det være lover og regler som endres i nær fremtid?
- c. Er virksomhetens drift, retningslinjer og prosedyrer i samsvar med loven?

Videre følger samling og analysering av informasjonen gruppen har hentet inn. Heretter merkes hvert punkt i forhold til viktigheten og potensiell risiko for nærliggende innvirkning for virksomheten din (CIPD, 2021). Videre kategoriseres disse ut ifra virksomhetens styrker og svakheter (SWOT-analyse) og prioriteres for å finne ut av hva virksomheten skal løse først. Identifiser så forretningsalternativene dine for å løse problemene. Ut ifra alternativene kan en handlingsplan utvikles for å maksimere styrker og redusere trusler.

Det vil videre være naturlig å lage en dokumentasjon eller sammendragsrapport for å dele internt i virksomheten, beslutningstakere og interessenter. Enkelte tiltak kan være kritiske og må tas stilling til fortløpende, andre kan overvåkes, noe handlingsplanen må ta høyde for. For å være effektiv, må en PESTEL-analyse gjøres regelmessig. Det er kun om en virksomhet har god kontroll på sine makroomgivelser, at de potensielt kan oppdage trender tidlig og gjøre dette til konkurransefortrinn.

#### 5.1.4 SWOT-analyse

En SWOT-analyse gjennomføres ved å først identifisere de interne styrkene og svakhetene. Med dette menes en evaluering av virksomhetens aktuelle ressurser, ferdigheter, ledelse, kultur, og andre mulige faktorer som kan ha en påvirkning på beslutninger. Dette innebærer Buitelaars åtte dimensjoner som en virksomhet består av. Videre identifiseres eksterne faktorer, gjennom trusler og muligheter, som kan være trender i markedet, handlinger gjort av konkurrenter, endring av regelverk, eller andre faktorer som kan ha en påvirkning på virksomhetens vedtatte mål og visjon. Når man skal analysere de forskjellige faktorene, vil det være aktuelt å stille seg følgende spørsmål (Sabbaghi & Vaidyanathan, 2004):

1. Styrker
  - a. Hva er våre fordeler?
  - b. Hva gjør vi bra?
  - c. Hvordan er vi det konkurransedyktig?
  - d. Hvilke ressurser har vi tilgjengelig?
  - e. Er det er noen interne ressurser (kunnskap, motivasjon, teknologi, finans, forretningsforbindelser) som vil bidra til møte krav og bekjempe trusler?
2. Svakheter:
  - a. Hva kan forbedres?
  - b. Hva er dårlig utført?
  - c. Hva bør unngås?
  - d. Er det noen interne underskudd/hindringer som hindrer virksomheten i å møte kravene?
3. Muligheter:
  - a. Hva er de gode oppgavene?
  - b. Hva er de interessante trendene?
  - c. Hvilke endringer forventer vi å se i markedet i løpet av det neste året?
  - d. Er det noen ytre omstendigheter eller trender som favoriserer etterspørselen etter en virksomhets spesifikke kompetanse?
4. Trusler:
  - a. Hva gjør konkurrentene våre?



- b. Hva er hindringene?
- c. Hvilke fremtidige endringer vil påvirke virksomheten vår?
- d. Er endring av teknologi en trussel for vår posisjon?
- e. Har vi ledelsesstøtte?
- f. Har vi tilstrekkelig med ressurser?
- g. Anvender vi de riktige verktøyene, programvarene og plattformene?
- h. Er det eventuelle ytre omstendigheter eller trender som vil påvirke etterspørselen etter en virksomhets kompetanse på en negativ måte?

Sett under ett, kan SWOT-analyse vise seg å være en verdifull del av strategisk planlegging for virksomheter. Analysen kan hjelpe virksomheter å evaluere sin nåværende situasjon, samt utvikle en strategi der man fokuserer på å maksimere styrker og minimere svakheter (Sabbaghi & Vaidyanathan, 2004). Videre kan man vurdere å bruke SWOT-analysen som en del av en større strategisk planleggingsprosess, der man kombinerer analysen med andre strategiske verktøy. Ved å foreta seg en slik kombinasjon vil man være i stand til å utvikle en veldefinert og kompleks strategi som tar hensyn til relevante faktorer ved virksomheten.

For å anvende en SWOT-analyse, har Christine Namugenyi et. al (2019) lagd en liste bestående av åtte steg for å lykkes med en systematisk SWOT-analyse. Disse åtte stegene er:

1. Velge mål for SWOT-analysen.
2. Undersøke virksomheten, bransjen og markedssituasjonen.
3. Identifisere og liste opp virksomhetens styrker.
4. Kategorisere og liste opp svakheter ved virksomheten og/eller bransjen.
5. Klassifisere og liste opp potensielle muligheter gjennom ulike perspektiver blant virksomheten og/eller bransjen.
6. Organisere og liste opp potensielle trusler for virksomheter som representert i bransjen.
7. Avgjøre prioriteringer fra innholdet i SWOT-analysen som er utført.
8. Utvikle, samt forbedre en strategi for å løse problemene og utfordringene som SWOT-analysen belyser, gjerne i flere bransjekontekster.

Ved gjennomførelsen av SWOT-analyser, er det viktig å primært vektlegge integreringen og analyseringen av de interne og eksterne faktorene ved en virksomhet eller bransje. Først og fremst vil det være essensielt å understreke formålet ved en slik analyse, da dette legger fundamentet for å holde virksomheten fokusert på SWOT-analysen, samt modelleringen av de ulike dimensjonene en bransje består av. Videre bør virksomheten gjøre forskning på tre forskjellige nivåer, virksomhets, bransje, -og markedsnivå. Dermed vil de få et komplekst syn på hvordan situasjonen er i alle de tre nivåene, og vil dermed opparbeide seg en forståelse de kan bygge analysen på. I denne fasen vil det være aktuelt å se sammenhenger av styrker, trusler, muligheter og trusler (Namugenyi et al., 2019). Avslutningsvis bør virksomheten være i stand til å utvikle prioriteringer basert på den foretatte analysen. Basert på informasjonen man sitter igjen med fra resultatet, vil man være handlingsdyktig til å utvikle en strategi, og senere implementere denne.

## 5.2 Dataforståelse

Den andre fasen i modellen er dataforståelse. Her er målet å undersøke hvilke data som er tilgjengelig for virksomheten, og gjøre vurderinger på hvilke data de skal benytte videre i prosessen. Basert på forståelsen etablert i forrige fase, skal virksomheter vurdere dette ut ifra behovene de har for å være i stand til å nå forretningsmålene. Virksomheten må vurdere både eksterne og interne faktorer. Interne faktorer kan være hvilke ressurser de har tilgjengelig, og eksterne kan være hvilke PESTEL-faktorer de vil undersøke gjennom analyse. De skal vite hvilke data som er tilgjengelige internt, samt hvilke forretnings spørsmål de er ute etter å besvare gjennom dataanalyse. Virksomheten skal sitte igjen med data utvalgt for videre analyse. Virksomheten skal ha en grundig forståelse av dataen i forhold til blant annet karakteristikker og kvalitet, samt hvordan bruk av denne dataen kan hjelpe virksomheten nå forretningsmål.

### Data

Basert på forretningsforståelsen må virksomheten her velge ut den dataen som best kan besvare spørsmålene virksomheten ønsker svar på. Dersom virksomheten ikke har tilstrekkelig data tilgjengelig internt, og heller ikke mulighet til å generere det, må dette hentes inn fra eksterne kilder. Dette er et av stedene hvor data-drevne makroanalyser skiller seg mest ut fra mikroanalyser. Data om makrofaktorer kan hentes inn fra enormt mange ulike kilder, så det vil være viktigere å være kjent med disse, samt sikre at valg blir basert på nøye vurderinger av hva virksomheten faktisk trenger. Selv om virksomheten sitter på store mengder data, er ikke dette alltid nok til å utføre gode analyser. Dette vil ofte være spesielt gjeldende for data-drevne makroanalyser. Grunnen til det er at det kan være relevant for flere virksomheter å benytte data om mange ulike makrofaktorer, men at dette ofte ikke er data som samles inn av virksomheten selv. Informant #2 nevner et konsept kalt «garbage in, garbage out»:

*«Hvis dataen er "garbage", vil dette også gjenspeiles i analysene de er i stand til å produsere. Sluttproduktene dataanalytikere er i stand til å produsere, gir lite verdi for virksomheten dersom det er mange problemer knyttet til datakvaliteten. Typiske problemer er inkonsekvens, nullverdier og ufullstendig data». - Informant #2*

*«For at bedrifter skal være kapable til å ta i bruk datadrevende metoder, kreves det i mange tilfeller data som allerede ikke er blitt samlet inn av virksomheten før. Man vil samtidig være nødt til å forholde seg til data som allerede eksisterer internt, samt at det foreligger tilstrekkelig grad av detalj som gjør at de kan brukes.» - Informant #3*

Virksomheten må gjøre seg kjent med de relevante eksterne kilder som er tilgjengelig, som kan benyttes til å besvare de etablerte forretnings spørsmålene. I tråd med CRISP-DM må dataen som hentes inn, også undersøkes i denne fasen. Avhengig av datakilden, kan dette være undersøkelse av hele datasett, et prøveutvalg direkte fra kilden, eller et prøveutvalg forespurt fra og levert av en dataleverandør (Legner et al., 2022).

Når virksomheten velger ut hvilke datakilder som skal integreres og benyttes, burde flere faktorer vurderes. Carl Anderson (2015, ss. 51–52) foreslår at data vurderes etter følgende dimensjoner:

- **Pris** – Det finnes mye data som er gratis å benytte. Likevel vil det noen ganger være bedre å betale for data av høyere kvalitet. Dersom data må kjøpes, må virksomheten vurdere om det er en rettferdig pris i henhold til verdien det vil gi virksomheten.

- **Kvalitet** – Dataen må vurderes etter hvor ren og pålitelig den er. Jo renere den er i utgangspunktet, jo mindre ressurser kreves for å behandle den slik at den kan benyttes. Dersom dataen ikke er pålitelig, vil den være mindre gunstig å benytte.
- **Eksklusivitet** – Er dataen som vurderes proprietær og eksklusiv? Dette kan ha stor innvirkning på hvilke fordeler det vil gi virksomheten ovenfor deres konkurrenter, dersom dataen benyttes godt.
- **Prøve** – Er det mulig å anskaffe et prøveutvalg av dataen, som vil tillate virksomheten å bedømme dataen etter kvalitet og innhold, samt gjennomgå formatet tilstrekkelig, uten å måtte inngå en fast kjøpsforpliktelse?
- **Oppdateringer** – Hvor ofte endres dataen? Blir den fort utdatert, eller irrelevant? Oppdateres den jevnlig? Dette kan ha mye å si for nytten dataen utgjør ovenfor virksomheten.
- **Pålitelighet** – Dersom dataen hentes gjennom en API; Hvordan er oppetiden? Er det noen begrensninger i forhold til API-kall eller tjenesteavtaler?
- **Sikkerhet** – Dersom det dreier seg om sensitiv data; er de kryptert og overført på en sikker måte?
- **Bruksvilkår** – Er det noen lisensvilkår eller begrensninger som kan hindre virksomheten i å hente ut verdi fra dataen?
- **Format** – Hvilket format kommer dataen i? Generelt foretrekkes det menneskelesbare -formater, som CSV, JSON eller XML, ettersom de ofte er enklere å tolke, og åpner opp for flere muligheter i å gjøre det på. Virksomheten må også vurdere om formatet allerede støttes av virksomhetens dataarkitektur, eller om de er nødt til å legge til ekstra utviklingskostnader og andre ressurser for å ta den i bruk.
- **Dokumentasjon** – Virksomheten bør foretrekke godt dokumenterte kilder. Det vil si - dokumentasjon på hvordan dataen er samlet inn, og hva som inngår i dataen av attributter og verdier, også kalt en dataordliste (eng. data dictionary).
- **Volum** – Kan virksomheten støtte volum- og lagringsbehovene til dataen? Det er ikke alltid slik at de største datasettene er de mest verdifulle. Geografisk data som f.eks. postnummer kan være svært verdifulle i datavitenskapsprosjektet tilknyttet markedsføring, men slike datasett er sjeldent store.
- **Granularitet** – Granularitet omhandler dybden av data. Går dataen ned på et tilfredsstillende informasjonsnivå?

Ved integrasjon av en ny datakilde burde virksomheten til beste evne foreta en nytte-kostnadsanalyse basert på blant annet de overnevnte faktorene. Det er viktig å vurdere om kostnadene ved å integrere datakilden i modellen som utvikles, overgår gevinsten (Anderson, 2015, s. 54). Alt dette må ses i perspektiv av forretnings spørsmålene som forsøkes å besvares.

Legner et al. (2022) foreslår at det basert på behovet for data, utarbeides et sett med vurderingskriterier for datakilde(n) som vurderes, enten de er interne eller eksterne. De identifiserer tre mulige varianter for hvordan vurderingskriteriene kan oppfylles. Variant A viser til at vurderingskriteriene er oppfylt, og at et nødvendig antall datasett er valg ut. Variant B viser til at ingen datasett (eller ikke en tilstrekkelig mengde med nødvendige datasett) har bestått vurderingskriteriene, og at vurderingskriteriene dermed bør revurderes. Variant C viser til at ingen datasett har bestått vurderingskriteriene selv etter revurdering, og at prosessen med å identifisere relevante datakilder burde revurderes. Denne vurderingsprosessen avsluttes med oppnåelse av en liste over valgte datasett som har bestått de etablerte vurderingskriteriene. Kvaliteten på ekstern data vil i stor grad

varierte, så fleksibiliteten denne prosessen fører med seg, vil være svært viktig. Hvis ingen datasett består vurderingskriteriene, må prosessen med identifisering av mulige datakilder gjennomgå på nytt for å identifisere ytterligere kandidatdatasett. Denne prosessen sørger for en strukturert tilnærming til vurdering av data.

Når datasettet eller datasettene er valgt ut, kan arbeidet med å integrere dem inn i virksomhetens egne systemer begynne. Dette er ofte en krevende oppgave, grunnet heterogeniteten til de utallige eksterne datasettene som er tilgjengelig. Basert på behovet etablert i fasen forretningsforståelse, samt en forståelse av virksomhetens egne interne integrasjonssystemer, burde virksomheten velge et passende grensesnitt for å få tilgang til utvalgt data. Dette kan være å laste ned og lagre dataen internt, eller å få tilgang til den gjennom et API. Det vil være viktig å grundig dokumentere de utvalgte datasettene. Dokumentasjonen burde gi fullstendig metadatainformasjon om datakilden og innholdet, mer spesifikt attributtene som inngår (Legner et al., 2022, s. 10).

## Metrikker

I fasen dataforståelse bør det utvikles metrikker med den hensikt å forstå dataen bedre, samt forstå hvor verdifull og brukbar den i praksis vil være for virksomheten. Dårlig datakvalitet kan medføre store kostnader for virksomheter, samt føre til at beslutningstakere ikke stoler på beslutningsgrunnlaget utviklet basert på data (Heinrich et al., 2018). Virksomheten bør dermed identifisere de dimensjoner av datakvalitet som vil være relevant for brukstilfeller av dataen de arbeider med, og utarbeide metrikker som muliggjør målinger av dem.

## Ferdigheter

I denne fasen vil det kreve ansatte med god kunnskap innen data. Data gir ikke nødvendigvis et realistisk bilde av virkeligheten, men må forstås av de ansatte som jobber tett med den. Det er de som jobber med dataen, som vil være i stand til å se muligheter og begrensninger ved den.

*«Sensoriske og IoT-data er gjerne ustrukturert og må tolkes og oversettes til data som "gir mening". Hva er det som gir mening og som er relevant? Prosessen med å oversette og forbedre datakvaliteten, blir gjerne gjennomført av mennesker. Menneskene som utfører denne oppgaven, må derfor ha kontroll og peiling på hva som kommer inn av informasjon, samt vite hva dataen gir informasjon om. De må forstå hva som er viktig, og hva som skal og kan brukes.» - (Informant #1)*

Det må sikres at de ansatte har relevante ferdigheter og kunnskap for å ta i bruk teknologien som trengs for å utvikle data-drevne makroanalyser. Dette innebærer å betrakte ansattes eksisterende kunnskap og ferdigheter i vurderinger av hvilken teknologi som skal implementeres, samt å tilby tilstrekkelig opplæring der det behøves. Ansatte burde tildeles roller i prosessen, basert på en grundig vurdering av hvor relevante ferdighetene deres er i forhold til de ulike rollene virksomheten etablerer.

## Teknologi

Virksomheten må teknologi som samsvarer med deres behov for data. I fasen dataforståelse omhandler dette teknologi for å understøtte innsamling-, lagring-, og databehandlingslaget samt innbefattede aktiviteter i virksomhetens dataarkitektur.

Nøyaktig hvilke teknologier som burde benyttes, avhenger av flere forretningsmessige faktorer og behov, både tekniske og organisatoriske. Det er vanlig å bruke ETL (Extract, Transform, and Load)-verktøy for å hente ut data fra forskjellige heterogene databaser, transformere dataene til et enhetlig standardformat, rens dem, og til slutt laste dem inn i virksomhetens systemer som f.eks. et datavarehus (Sreemathy et al., 2021). Forskjellige verktøy har ulike styrker og svakheter. Det finnes ikke et universalt ETL-verktøy som passer alle virksomheter og deres interne behov (Kherdekar & Metkewar, 2016).

Eksempler på ETL-verktøy er IBM Infosphere DataStage (Sreemathy et al., 2021). Fordeler ved dette er at det er fleksibelt og historisk høyt anerkjent. Ulemper er at det er vanskelig å lære, krever lang tid å implementere, og at det krever en høy mengde arbeidsminne og prosessorkraft for å benytte. Et annet eksempel er Microsoft SQL Server Integration Services (SSIS). Fordeler ved dette er at det er enkelt å implementere, har lav kostnad sammenlignet med andre alternativer, og at leverandøren tilbyr god kundestøtte. Ulemper er at det blant annet er begrenset til Windows, som gjør det vanskelig å ta i bruk om virksomheten din hovedsakelig baserer seg på et annet operativsystem som f.eks. Linux (Sreemathy et al., 2021).

Valg av teknologi burde vurderes ut ifra den etablerte forretningsforståelsen og faktorer sentrale for virksomheten. Faktorer som kan være nyttig å ta i betraktning er blant annet brukervennlighet, skalerbarhet, sikkerhet, kundesupport, stabilitet, om virksomheten skal benytte det til batch- eller sanntidsdata, pris, og kompatibilitet med eksisterende teknologi (Saras, 2022).

## Ledelse

Eksisterende litteratur tyder på at det vil være essensielt å etablere dedikerte organisatoriske roller med ulike ansvarsområder, langs prosessen med å innhente og forvalte ekstern data (Belissent, 2019; Aaser & McElhaney, 2021; Explorium, 2021; som sitert i Legner et al., 2022). Mange av aktivitetene i denne prosessen kan allerede utfylles av tradisjonelt eksisterende roller. Tross dette tyder den begrensede litteraturen på at dersom innhenting av ekstern data er nødvendig for at virksomheten skal være i stand til å utvikle makroanalyser, er dataspeider (eng. data scout) en fremvoksende og potensielt essensiell ny rolle (Aaser & McElhaney, 2021; Legner et al., 2022). Denne rollens ansvar vil være å finne, gjennomgå, vurdere, og lede forvaltningen av eksterne data-ressurser (Aaser & McElhaney, 2021). Rollen vil opparbeide seg ekspertise på området, og fungere som et internt kontaktpunkt i henhold til ekstern data.

Legner et al. (2022, s. 11) foreslår følgende roller og tilknyttet ansvar i prosessen med å ta inn ekstern data, oppsummert i tabellen nedenfor:

**Tabell 4 Roller i dataforståelse**

| <b>Rolle</b>                                      | <b>Ansvar over aktiviteter i prosessen</b>   |
|---|--|
| <b>Forespører</b> (eng. requestor)                | <ul style="list-style-type: none"> <li>- Sender inn en forespørsel etter ekstern data.</li> <li>- Benytter dataen i utvalgte brukstilfeller.</li> </ul>  |
| <b>Dataanalytiker og forretningsanalytiker</b>    | <ul style="list-style-type: none"> <li>- Definerer krav til forretning og data for fremtidige brukstilfeller.</li> <li>- Identifiserer nye brukstilfeller for ekstern data, utvikler gjennomførbarhetsbevis, og gjennomfører analyser.</li> </ul>  |
| <b>Dataarkitekt</b>                               | <ul style="list-style-type: none"> <li>- Definerer systemkrav for integrasjon av ekstern data</li> </ul>   |
| <b>Dataspeider</b>                                | <ul style="list-style-type: none"> <li>- Identifiserer, vurderer og forvalter ekstern data.</li> <li>- Overvåker eksterne datakilder for relevant data som kan muliggjøre brukstilfeller.</li> <li>- Fungerer som et internt kontaktpunkt i virksomheten for å innhente og forvalte ekstern data.</li> <li>- Gir førstelinjes støtte internt, og fungerer som intern konsulent for emner relatert til ekstern data.</li> </ul> |
| <b>Dataansvarlig</b> (eng. data steward)          | <ul style="list-style-type: none"> <li>- Vurderer kvaliteten og kompatibiliteten av eksterne datasett med forretningen.</li> </ul>   |
| <b>Innkjøpsansvarlig</b> (eng. procurement)       | <ul style="list-style-type: none"> <li>- Forhandler kontrakter og tjenesteavtaler med dataleverandører, analyserer priser, sier opp kontrakter.</li> </ul>   |
| <b>Samvarsansvarlig</b> (eng. compliance officer) | <ul style="list-style-type: none"> <li>- Sikrer at eksterne data samles inn med egnede tillatelser (kontrakt- og lisensavtaler) og brukes i samsvar med gjeldende lover om data og interne retningslinjer.</li> </ul>  |

## Kultur

Ettersom data-drevne makroanalyser er et relativt nytt fagområde, eksisterer det ingen spesielt veletablerte prosesser eller metoder for hvordan virksomheter burde utvikle dem. Bruk av ekstern data åpner i tillegg opp et enormt antall nye muligheter, hvor virksomheter selv må utforske hvordan de skal gå fram, uavhengig av eksisterende kunnskap. Virksomheter som ønsker å ta i bruk data-drevne makroanalyser, vil i stor grad stå på egne ben i denne fasen. Data-drevne makroanalyser kan potensielt føre til store konkurransefortrinn for de virksomheter som er i stand til å benytte dem effektivt, men det vil kreve en kultur som er åpen for endring og innovasjon, og villig til å prøve ut nye ideer og tilnærminger. Veien mot innovasjon består ofte av mye prøving og feiling, noe som kan være spesielt gjeldende i denne fasen. Det vil være viktig å skape en innovasjonsfremmende kultur, for å få mest ut av et lite utforsket område som data-drevne makroanalyser. Et kjennetegn på en god kultur som støtter virksomheters ønske om å bli mer data-drevne, er en vilje til å eksperimentere og teste nye idéer (Anderson, 2015., som sitert i Berndtsson et al., 2020).

For å skape en innovasjonsfremmende kultur, burde virksomheten blant annet være åpen for å prøve ut nye ting, uten at ideer som mislykkes straffes (P. K. Ahmed, 1998). Virksomheten burde heller forsøke å forstå hvorfor ideer mislyktes, og deretter ta læring fra dette og benytte lærdommen videre. Virksomhetskulturen burde fremme tillitt og åpenhet, derunder god kommunikasjon og samarbeidsvilje. Nye ideer burde bli oppfordret og verdsatt, både av andre ansatte og ledelsen. Det vil være viktig å sørge for at innovative ansatte føler seg hørt, og at oppnåelser blir belønnet. I stedet for at ledelsen påtar seg det overordnede ansvaret for innovasjon, kan de tjene mye på å heller fokusere på å skape en kultur og et miljø som fremmer innovasjon blant de ansatte (P. K. Ahmed, 1998). Spesielt med tanke på data-drevne makroanalyser, vil det ofte være slik at de underordnede ansatte som jobber nærme dataen, er de som forstår den best. De kan potensielt sitte på mye verdifull kunnskap ledelsen ikke innehar, knyttet til begrensninger og muligheter for å benytte data til makroanalyser. For å utnytte dette, vil det være viktig å skape en kultur som motiverer de ansatte til å forsøke å benytte denne kunnskapen.

## Strategi

Det vil være viktig med en tydelig definert og kommunisert strategi. Dette skal sørge for at de ansatte involvert i arbeidet med å utvikle makroanalyse, er klar over hva virksomheten ønsker å oppnå. I denne fasen vil dette gi en forståelse over hva som kreves av dataen(e) som skal benyttes. Uten dette, vil det være vanskelig å foreta seg vurderinger om hva som utgjør den riktige dataen fra et forretningsperspektiv.

*«Hva som kreves av data, avhenger av hva virksomheten ønsker å oppnå. Ulike virksomheter har ulike behov, avhengig av deres strategi og mål» - (Informant #1)*

Når det gjøres vurderinger ovenfor hvilke datakilder virksomheten skal integrere, er det dermed viktig at de ansatte som er ansvarlig for aktiviteter her, er underforstått virksomhetens strategi. Uten at de er det, vil de ikke klare å handle på måter som best mulig gagnar virksomheten og dens mål. Klare mål burde utvikles for prosjektet, basert på virksomhetens overordnede strategi.

## Smidighet

*«Den største utfordringen knyttet til smidighet er å holde koblinger oppdatert og fange opp når f.eks. datakilder endres eller slutter å svare. Løpende vedlikehold og overvåkning er dermed veldig viktig siden slike systemer gjerne har mange komponenter og er sårbare for endringer» - (Informant #4)*

En virksomhets IT-arkitektur består ofte av mange sammenhengende komponenter som er sårbare for endring. For å sikre smidighet kreves det at IT-arkitekturen holdes oppdatert og relevant. Gjennom dette vil man øke sannsynligheten for at virksomheten forholder seg relativt kompatibel med eventuelle endringer av krav til datakilder og verktøy. Virksomheter burde vurdere skyløsninger der det er mulig, da slike løsninger ofte er mer fleksible i forhold til blant annet skalerbarhet.

Dersom ekstern data blir integrert i virksomhetens interne systemer, vil det fortsatt være kritisk å sikre at den blir brukt på en god måte. Det burde tas hensyn til aspekter ved ekstern data som er ulik intern data. Endringer utenfor virksomhetens kontroll, kan

forekomme i de eksterne datakildene. Dette kan være endringer i innholdet, eller at datakilden rett og slett slutter å svare, som beskrevet av informant #4. Virksomheter som tar i bruk ekstern data, burde sørge for å etablere gode prosesser for å overvåke og forvalte de integrerte datakildene hvor slike endringer kan forekomme. Dersom virksomheten benytter eksterne datakilder som oppdateres jevnlig, burde dette håndteres gjennom etablering av versjoner av datasett fra datakilden. Nye versjoner av datasettet, burde integreres gjennom etablerte prosesser for nye kilder. Virksomheten burde i tillegg sørge for at tilbakemeldinger om dataen fra brukere blir dokumentert. Dette kan sikre bedre datakvalitet, og identifisere muligheter for forbedring (Legner et al., 2022).

### 5.3 Dataforberedelse

Dataforberedelse er den tredje fasen. Fasen starter med den utvalgte dataen videreført fra forrige fase. Fasen tar for seg alle aktiviteter som kreves for å konstruere det endelige datasettet som skal benyttes til å utvikle modeller i modelleringsfasen. Avhengig av brukstilfellet, må dataen renses og transformeres for å møte virksomhetens behov for analyse.

#### Data

Basert på de foregående fasene skal virksomheten til beste evne vite hva som kreves av data for å oppnå forretningsmålene. De skal også ha vurdert og valgt ut datasett for videre bruk, samt vurdert datakvaliteten mot tenkt brukstilfelle(r). Konseptet «garbage in, garbage out», er kritisk også her. Det første som burde gjøres, er å velge ut den dataen som er relevant for brukstilfellet. Virksomheten vil ikke alltid behøve all dataen i et datasett, men bør tidlig velge hva som skal inkluderes og hva som skal ekskluderes basert på forretningsmålene. Dette kan for eksempel være et utvalg av rader eller kolonner i et eller flere datasett. Dette kan være spesielt viktig i henhold til bruk av ekstern data for makroanalyser, ettersom mye av den tilgjengelige dataen finnes i store åpne datasett hvor kun deler av dataen vil være relevant for det definerte brukstilfellet.

Når dataen som skal benyttes er valgt ut, må dataen renses for å øke datakvaliteten. Dette er et kritisk steg. Urenset data kan inneholde feil som fører til et urealistisk beslutningsgrunnlag dersom det ikke blir fikset. Rensing av data består av to hovedsteg. Det første er å oppdage feil, og det andre er å reparere feilen. For å oppdage feil, burde virksomheter ifølge Chu et al. (2016), basere sine valg av teknikker på tre spørsmål. Disse er; «Hva skal oppdages?», «Hvordan kan det oppdages?», og til slutt «Hvor skal det oppdages?». For å gjøre tilsvarende i forhold til hvordan feilene skal repareres, sier de at virksomheten burde stille seg spørsmålene; «Hva skal repareres?», «Hvordan skal det repareres?» og «Hvor skal det repareres?»

#### **Oppdagelse av feil**

**Hva skal oppdages?** – Virksomheten må være klar over mulige feil i data, samt hva konsekvensene av dem kan være. Typiske feil kan være avvikende ekstremverdier, manglende verdier, duplikater, formateringsfeil, etc. (Chu et al., 2016; Talend, u.å.)

*«Sluttproduktene de er i stand til å produsere gir lite verdi for virksomheten dersom det er mange problemer knyttet til datakvaliteten. Typiske problemer er inkonsekvens,*



*nullverdier og ufullstendig data. Løsningene som er tilgjengelig, må drøftes og diskuteres med kunden. Hvis man ikke vet hva man ønsker å bruke produktet til, så er det vanskelig å vite hva man trenger» - (Informant #2)*

**Hvordan skal det oppdages?** – Avhengig av type feil i data, kan ulike teknikker benyttes for å oppdage dem. Mange teknikker kan fullstendig automatiseres, for eksempel oppdagelse av overtredelser av funksjonelle avhengigheter i en database (Bohannon et al., 2005). Andre teknikker igjen, som oppdagelse av duplikater, involverer ofte mennesker i undersøkelsesprosessen (Wang et al., 2012).

**Hvor skal det oppdages?** – Feil kan oppstå i alle trinn i en dataarkitektur og datavitenskapsprosess. For eksempel blir feil i kilden til databasen ofte spredt gjennom dataprosessering. Mens de fleste feildeteksjonsteknikker oppdager feil i den opprinnelige databasen, kan noen feil bare oppdages mye senere i data-prosesseringen (Chu et al., 2016). Virksomheten må være klar over hvor typiske feil i dataen de skal benytte oppstår, og hvor det så kan oppdages.

### **Reparering av feil**

**Hva skal repareres?** – Det er flere måter å reparere feil i data på. Dataen kan for eksempel repareres direkte i kilden, ved at mennesker eller programmer gjør direkte endringer i datasett. Det kan også kreve endringer i begrensninger i databasene, for å sikre at dataen blir lagret korrekt (Chu et al., 2016). I følge blant annet informant #1, kan feil i data ofte også skyldes menneskelige handlinger.

*«Data blir ofte ikke bedre enn input den får. Mye data blir lagt inn manuelt av mennesker, hvor feil kan oppstå grunnet svake prosesser og rutiner. Datakvaliteten må være så nøyaktig og god som mulig.» - (Informant #1)*

Det vil derfor være viktig å avgjøre hva som skal repareres. Virksomheten burde stille seg spørsmål knyttet til om det for eksempel holder å reparere dataen direkte, om det krever endringer i virksomhetens databaser, eller om feil i internt generert data skyldes dårlige rutiner og prosesser, som igjen fører til dårlig menneskelig input og derav feil i data.

**Hvordan skal det repareres?** – Det eksisterer ulike metoder for å reparere feil i data. Noen feilkilder kan fikses fullstendig automatisk, mens andre krever at mennesker involveres i prosessen for å verifisere reparasjoner, foreslå mulige reparasjoner, eller for å trene opp maskinlæringsmodeller for å automatisere reparasjonsprosesser (Chu et al., 2016; Yakout et al., 2011). Virksomheter må vurdere hvilke teknikker de skal benytte ut ifra blant annet type feil de må reparere, og virksomhetens eksisterende ressurser.

**Hvor skal det repareres?** – De fleste vanlige teknikker reparerer feil i databaser ved å gjøre direkte endringer i den. Dette endrer dermed den originale databasen. For ikke å gjøre dette, bygges det ofte en modell for å beskrive de forskjellige måtene å reparere den underliggende databasen på. Spøringer besvares av disse modellene ved å bruke for eksempel et utvalg fra alle mulige reparasjoner (Beskales et al., 2010; Chu et al., 2016).

Når prosessen med å rense data er over, må dataen valideres ved å teste for feil i dataforberedelsesprosessen opp til dette punktet. Ofte vil en feil i systemet bli tydelig under denne valideringsprosessen. Dette må løses før man går videre (Talend, u.å.).

Virksomheten må være svært nøye med denne prosessen, for å sikre at alle mulige feil blir oppdaget, og deretter korrigerert på best mulig måte. Dette må gjøres for å unngå de store negative konsekvensene dårlig datakvalitet kan ha på en analyse, og de videre beslutningene befunnet på analysen.

Videre vil det ofte være hensiktsmessig å transformere dataen. Dette vil si å oppdatere formatet eller verdiene i oppføringene for å oppnå et veldefinert resultat, eller å gjøre dataene lettere forståelig for et bredere publikum (Talend, u.å.). Ifølge informant #1, er standardisering av data også viktig å tenke på her:

*«Jo renere dataen er i utgangspunktet, jo enklere er den å behandle. Grad av standardisering er et viktig element her. Dersom dataen er likere, vil datakvaliteten oppleves høyere. Standardisering kan være å for eksempel alltid sørge for at samme måleenhet blir brukt når data registreres» - (Informant #1)*

Det vil i tillegg ofte være nødvendig å berike data. Dette refererer til å legge til og koble data med annen relatert informasjon for å gi dypere innsikt. Dette kan være å integrere ulike datasett med hverandre, dersom man for eksempel har data fra flere kilder som skal benyttes til å besvare de samme forretnings spørsmålene (IBM, 2021d; Talend, u.å.). Dette kan også være spesielt viktig i henhold til data-drevne makroanalyser. Virksomheter kan potensielt kombinere egen intern data med ekstern data, og kontekstualisere den interne dataen med data om makroomgivelser.

Etter at alle aktivitetene i fasen dataforberedelse er gjennomført, skal virksomheten sitte igjen med det endelige datasettet som videre skal benyttes i neste fase (IBM, 2021d).

## Metrikker

Det vil være viktig å benytte hensiktsmessige metrikker i denne fasen av et dataanalyseprosjekt. Her vil det også være viktig å ha forretningsmålene. Noen brukstilfeller kan kreve svært høy datakvalitet målt etter spesifikke dimensjoner som for eksempel nøyaktighet, mens det for andre ikke vil være like kritisk at absolutt all dataen er nøyaktig. Styring av datakvalitet må støttes av forretningsmessige begrunnelser. Datakvalitet kan fikses hele dagen lang, men det betyr ikke nødvendigvis at det kommer virksomheten til gode (Turner, u.å., som sitert i Stedman, 2020). Virksomheter bør dermed ikke bruke sett med standardiserte metrikker for all dataen de forbereder. De bør heller utvikle og tilpasse metrikker basert på hva dataen de måler skal benyttes til.

Stedman (2020) presenterer noen typiske metrikker virksomheter kan benytte:

1. Statistikk over antall datafeil som blir identifisert og rettet per måned eller kvartal.
2. Nøyaktighet og feilrate i datasettene, med varsler hvis feil overstiger akseptable nivåer.
3. Kvantitative mål etter dimensjoner av datakvalitet som er viktige for virksomheten.
4. Beregninger på virkningen av datakvalitetsproblemer og potensielle reparasjoner på virksomheten.
5. Undersøkellesdata og annen tilbakemelding om datakvalitet, samlet inn fra dataens sluttbrukere.

## Ferdigheter

Det er viktig at de ansatte ansvarlige for datakvalitet, innehar relevant kunnskap innen det tekniske som kreves for å sikre den. Det vil også være viktig at de innehar tilstrekkelig kunnskap om det organisatoriske, for å kunne sikre at arbeidet deres gir verdi fra et forretningsmessig perspektiv. I tillegg burde det vektlegges å opparbeide en generell datakompetanse innad virksomheten. Dette innebærer at også de ansatte som benytter seg av data i sitt arbeid, uten å overordnet være ansvarlig for datakvaliteten, er i stand til å forstå data og se begrensninger og muligheter ved den (Korolov, 2020). Dette kan gjøre virksomheter med effektive i sitt arbeid med å sikre god datakvalitet for gitte brukstilfeller. Hvordan dette kan gjøres diskuteres under kultur og ledelse i samme fase nedenfor.

## Teknologi

Ulike ETL-verktøy innehar forskjellige metoder for å forberede dataen. Dersom virksomheten ikke benytter seg av ETL-verktøy, eller at det valgte verktøyet ikke dekker funksjoner for nødvendige aktiviteter i denne prosessen, kan virksomheten supplementere med spesialiserte verktøy for dataforberedelse. Det finnes for eksempel verktøy som kan brukes til å automatisere oppgaver knyttet til datarengjøring, inkludert både kommersiell programvare og åpen kildekode teknologier. Typisk inneholder verktøyene en rekke funksjoner for å korrigere datafeil og -problemer. Mange utfører også data-sammenligning for å finne duplikater (Stedman, 2022).

## Ledelse

Arbeidet rundt forberedelsen av data, blir vanligvis utført av spesialiserte datakvalitetsanalytikere eller andre fagpersoner innen data. Det er også vanlig at dataanalytikere og andre brukere av data, bistår i prosessen med å forberede dataen til eget bruk (Stedman, 2022).

For å sikre at involverte ansatte innehar relevante ferdigheter og tilhørende datakompetanse for å benytte data i arbeidet sitt, krever det en ledelse som aktivt bistår i arbeidet med å lære opp de ansatte.

For øyeblikket eksisterer det ingen beste praksiser virksomheter kan følge når det gjelder å lære opp ansatte i datakompetanse. Det finnes derimot flere ulike tilnærminger som virksomheter kan vurdere (Korolov, 2020). Én tilnærming er å tilpasse utdanning til de ulike rollene ansatte har i en virksomhet og arbeidet de utfører. Dette kan for eksempel gjøres gjennom opplæringsprogram som Data Science Academy tilbudt av Coursera. Virksomheter kan også gi opplæring til folk når de trenger det, for eksempel når de trenger å benytte et nytt dataverktøy. En annen tilnærming er å tilby målrettede, praktiske workshops til ansatte med fokus på de spesifikke analyseverktøyene som brukes i selskapet og de spesifikke forretningsutfordringene rundt data som selskapet og deres ansatte står overfor. Disse workshopene kan arrangeres enten av verktøyleverandørene selv, eller av uavhengige konsulenter (Korolov, 2020). Det vil være viktig at de ansatte og deres ledere forstår fordelene med opplæringsprogrammene. De bør ha klare svar på spørsmål som "hva får jeg ut av dette?" og "hvordan er opplæringen relatert til min nåværende eller fremtidige rolle?" (Panetta, 2021).

## Kultur

Det vil være viktig å bygge opp en kultur hvor data blir verdsatt av virksomhetens ansatte. Data- og analyseledere bør være ansvarlige for å fremheve den forretningsmessige verdien som kan oppnås gjennom bruk av data (Panetta, 2021). Kulturen burde fremme datakompetanse, hvor datakvalitet sees på som essensielt. Virksomheter kan tjene mye på at alle involverte ansatte er i stand til å bidra i prosessen med dataforberedelse til deres brukstilfeller.

Amirah Fayek foreslår at virksomheter utvikler metrikker for å bistå i arbeidet med å opparbeide en kultur hvor datakvalitet verdsettes (Fayek, u.å., sitert i Stedman, 2020). Eksempler på slike metrikker sier hun kan være:

1. Antall henvendelser om data fra forretningsbrukere.
2. Antall unike pålogginger fra ansatte i en datainnsiktsportal.
3. Prosentandelen av ansatte som deltar på arrangementer rundt opplæring i datakompetanse.
4. Prosentandelen av ansatte som rapporterer om økt kunnskap etter opplæringsøkter og arrangementer.
5. Prosentandelen av databrukere som sier at deres data er i tråd med virksomhetens målsetninger.

## Strategi

Behov for strategi i denne fasen tilsvarer behovet i fasen dataforståelse, dvs. at det vil være viktig med en tydelig definert og kommunisert strategi. Ulike mål krever varierende data. En tydelig strategi med klare mål gjør det mulig å utarbeide tydelige mål for analyseprosjektet, som dermed vil sørge for at dataen blir forberedt etter disse.

## Smidighet

I fasen dataforberedelse vil det være krav til smidige prosesser. Det kan oppstå problemer underveis i prosessen, eller det kan alternativt dukke opp muligheter som tidligere ikke har blitt vurdert. Dataforståelsen kan endre seg i løpet av dataforberedelsesfasen, og det kan vise seg nødvendig å tilpasse dataen eller metodene for å forberede den, for å videre kunne bruke den til analyser. Her vil det også være essensielt å ha en oppdatert og relevant IT- og dataarkitektur, for å øke sannsynligheten for at virksomheten forholder seg kompatibel med eventuelle endringer av krav til datakilder og verktøy.

## 5.4 Modellering

I modelleringsfasen bygger, og vurderer teamet i virksomheten diverse modeller opp mot prosjektets bestemte tematikk. Dette skjer gjerne gjennom en rekke forskjellige modelleringsteknikker, der teknikken defineres av virksomhetens begrensninger. Hovedformålet med denne fasen er å velge modelleringsteknikker, generere testdesign, bygge, og vurdere modellene (J. S. Saltz, 2021).

## Data

Basert på tidligere faser skal dataen være forstått og forberedt i henhold til ønsket formål basert på et reelt forretningsbehov. I modelleringsfasen er det sentralt å ta høyde for hvilke data modellen skal representere i henhold til de data-drevne metodene. For å lykkes med modelleringen er det dermed essensielt å se de data-drevne makroanalysene i henhold til den tilgjengelige dataen. Dermed forsikrer man seg at resultatet blir så nøyaktig som mulig, i henhold til det ønskede resultatet (Wirth & Hipp, 2000). Dette gjør at man opparbeider seg det beste utgangspunktet for å velge den mest passende modellteknikken til utarbeidelsen av de data-drevne makroanalysene. Det eksisterer enormt mange ulike verktøy og modelleringsteknikker, og hva som utgjør riktig valg avhenger av både formålet og dataen i seg selv.

Virksomheten må gjøre seg kjent med de ulike verktøyene og modelleringsteknikkene som kan være relevante for dem og tenkt brukstilfelle. Videre foreslår IBM (2021h) å stille følgende spørsmål rundt valg av modelleringsteknikk:

1. Kreves modellen at dataene deles inn i test- og treningssett?
2. Har du nok data til å produsere pålitelige resultater for en gitt modell?
3. Krever modellen et bestemt nivå av datakvalitet? Kan du oppnå dette nivået med de nåværende dataene?
4. Er dataene dine av riktig type for en bestemt modell? Hvis ikke, kan du gjøre de nødvendige konverteringene slik at de blir det?

Modellen må deretter bygges basert på valgt modelleringsteknikk.

## Metrikker

Modelleringsfasen må primært ta for seg hvilke metrikker som er hensiktsmessig for å måle modelleringsfasens effektivitet. Metrikkene kan dermed ta for seg ting som ressursbruk, tiden modelleringen tar, etc. (Wirth & Hipp, 2000). Ved å gjøre relevante målinger, vil man tilegne seg det beste grunnlaget til å foreta en vurdering av de aktuelle modellene og prosessen rundt, og kan dermed gjøre et reflektert valg. Dersom målinger f.eks. viser at mye tid og ressurser blir brukt på modeller som ikke viste seg brukbare, kan dette tyde på problemer knyttet til alle foregående faser som f.eks. dårlig forretnings-/dataforståelse.

## Ferdigheter

Riktige ferdigheter er essensielt i implementeringsfasen. At de ansatte kan anvende de aktuelle teknologiene og verktøyene, samt har kompetanse innenfor data-drevne makroanalyser vil være nøkkelen til suksess. For at man skal lykkes med modelleringen av de data-drevne makroanalysene, er man nødt til å ha kompetanse rundt felt som å tolke resultater, skissere og bygge modeller, samt foreta justeringer på diverse parameterinnstillinger blant den aktuelle teknologien.

## Teknologi

I modelleringsfasen er det essensielt at virksomheten besitter aktuell teknologi for å kunne gjennomføre modelleringen suksessfullt (Wirth & Hipp, 2000). Dermed vil man også måtte ha en forståelse om hva de bør skaffe seg, eller oppgradere for å være i stand til å gjennomføre modelleringen. Samtidig er virksomheten nødt til å ha teknologi

som er handlingsdyktige nok til å kunne generere testresultatene, teste og validere disse, og deretter hente ut verdien fra disse resultatene.

## Ledelse

Ledelsen vil være ansvarlig for at de som utvikler modellen har forretningsforståelse, og kompetanse innenfor data-drevne makroanalyser, slik at modelleringen står i henhold til disse faktorene. Dette gjøres for å sikre seg optimale resultater. Samtidig er man avhengig av at ledelsen overser og forsikrer at modelleringen skjer i samsvar med strategi, og formål med den data-drevne makroanalysen. Dette gjøres for å korrigere at faktorene samsvarer, og at analysene skal gi forretningsverdi.

## Kultur

Dette dekkes i stor grad av kultur under fasen dataforberedelse. I tillegg vil det her være viktig med ansatte som stiller seg kritisk til resultatene modelleringsfasene tilføyer. Dette kan være meget sentralt for å oppnå best mulige resultater, og for å korrigere eventuelle feil (Liu et al., 2021). Kulturen i modelleringsfasen bør derfor innebære å stille seg kritiske til resultatene, samt være nøye på at modelleringen skjer i samsvar med virksomhetens strategi. Samtidig bør man være reflekterte når man foretar seg evalueringene av modellresultatene, slik at man har et veldefinert grunnlag til å ta en avgjørelse.

## Strategi

Behov for strategi i denne fasen tilsvarer behovet i fasen dataforståelse og dataforberedelse, dvs. at det vil være viktig med en tydelig definert og kommunisert strategi. Ulike mål krever ulike modeller. En tydelig strategi med klare mål gjør det mulig å utarbeide tydelige mål for analyseprosjektet. Ved å etablere en strategi, sørger man også for at man har bestemte rammer å forholde seg til, og at modelleringen skjer innenfor disse.

## Smidighet

I modelleringsfasen bør man utvikle nøyaktige modellbeskrivelser basert på resultatene (Wirth & Hipp, 2000). Ved å dokumentere resultatene, vil man til enhver tid har oversikt over hva som er gjort. Dette gjør at virksomheten raskt skal kunne tilpasse seg endringer i data, datakilder, hyppighet, oppdateringer og kvalitet i dataen. Modelleringsfasen og dataforberedelsesfasen henger ofte tett sammen. I modelleringsfasen vil man ofte avdekke problemer med dataen, eller eventuelt få nye ideer for hva som kan konstrueres av data i foregående fase (Wirth & Hipp, 2000). Dermed vil det ofte være nødvendig eller gunstig å gå tilbake til en av de foregående fasene, dersom dette gjør at forretnings spørsmålene blir bedre besvart.

## 5.5 Evaluering

Som nest siste fase i CRISP-DM, finner vi evalueringsfasen, der man evaluerer de foregående modelleringene. Modellene evalueres etter om de skal implementeres eller ikke. Videre omhandler fasen å kontrollere at resultatene er i henhold til de definerte

forretningsmålene (J. S. Saltz, 2021). Derfor er resultatene nødt til å tolkes, samtidig som ytterligere handlinger må defineres. Avslutningsvis bør man forsikre seg om at hele prosessen blir gjennomgått og vurdert.

## Data

Et sentralt aspekt i evalueringsfasen er å vurdere dataene som anvendes i modellen. Data er fundamentet for å kunne gjennomføre data-drevne metoder, og det er derfor meget viktig å kritisk vurdere om man har tilstrekkelig med data, og om dataen er forsvarlige (Chapman, 1999). Dermed vil man være avhengig av å vurdere datakvaliteten opp mot bruksmålet for dataen. En gjentakende problemstilling er at virksomheter ikke har god nok datakvalitet (J. S. Saltz, 2021). Dermed vil det være aktuelt å foreta seg en evaluering av datasettenes egnethet og tilgjengelighet, og gjerne vurdere dette i henhold til fremtidig bruk. Et eksempel på en måte å vurdere datakvalitet kan være gjennom fenomenet «wall of shame».

*«"Wall of shame" er et typisk fenomen som tas i bruk for å sikre god datakvalitet blant bedriftene. Dette innebærer å markere hvilke data som er mangelfull, eller hvilke prosesser eller ansatte som er mangelfulle. Dermed vil man være rustet til å kunne skille ut avvik, og hvor i prosessen det oppstår avvik.» - (Informant #3)*

Gjennom å markere generelle avvik i dataene og prosessene, gjør dette at man sikrer seg at den innsamlede dataen som analysene skal bygge på, er forsvarlige til bruk, og at den deretter gir tilfredsstillende resultater.

## Metrikker

Evalueringsfasen bør fokusere på å evaluere modellresultatene opp mot forretningsmålene som er satt for prosjektet, altså å gjøre virksomheten i stand til å anvende data-drevne makroanalyser. For å kunne evaluere modellresultatene opp mot forretningsmålene, er man nødt til å foreta seg evalueringer av enkelte faktorer hos modellen. Disse faktorene utgjør primært hvor presis modellen i forhold til bruksformål, og dens kapasitet til å predikere potensielle hendelser eller forklare ønskede mønstre (Ayele, 2020).

## Ferdigheter

Innad i virksomheten bør man besitte ferdigheter som kan evaluere og teste dataene som modellen utgjør, opp mot virksomhetens bestemte forretningsmål (Wirth & Hipp, 2000). En evaluering og test av modellen opp mot virksomhetens forretningsmål innebærer å sjekke om modellen kolliderer med eventuelle hindringer, eller problemer, i henhold til forretningsmålene (Heggernes, 2020, s. 19–23). Dette gjøres gjennom en prosess der man stegvis går gjennom aktuelle faser i modellen, for å deretter vurdere disse kritisk. Dermed tilegner man seg også en forståelse av hvorvidt man bør foreta seg en ny evaluering av modellen, om modellen er levedyktig, eller om det er enkelte områder som bør revideres. Som et resultat av denne evalueringen, forsikrer man seg en trygg evaluering, der sannsynligheten for å lykkes med data-drevne makroanalyser øker (Chapman, 1999). For å kunne foreta seg slike evalueringer, er det essensielt at virksomheten besitter ansatte med de riktige ferdighetene for å støtte evalueringsprosessen (Wirth & Hipp, 2000). I det nevnte tilfellet, vil ferdighetene primært utgjøre kompetanse blant de ansatte. Denne kompetansen er nødt til å fokusere

på selve utviklingen til modellen. Som et resultat av denne gjennomgangen vil man være i bedre stand til å evaluere styrker og svakheter ved CRISP-DM-prosessen, samt å identifisere forbedringsområder ved aktuelle prosjekter. For ledelsen er dette til stor hjelp, spesielt når de skal planlegge virksomhetens videre arbeid med data-drevne makroanalyser.

## Teknologi

Basert på det teknologiske, bør det foretas en evaluering av virksomhetens infrastruktur. Det er derfor aktuelt å stille seg spørsmål om virksomhetens infrastruktur, slik at man forsikrer seg at den er i stand til å håndtere det data-drevne makroanalyser krever. For å evaluere dette, vil man typisk fokusere på egenskaper blant virksomhetens maskinvare, programvare, nettverk og operativsystem. Dersom man foretar seg evaluering av disse komponentene, vil man forstå om virksomheten yter maksimal kapasitet i prosessen rundt utvikling av data-drevne makroanalyser. Dersom den ikke gjør det, kan man gjøre tester på disse komponentene, slik at man finner ut hvor man bør foreta seg justeringer. Ved å gjøre slike tester, vil man luke ut svakheter i virksomhetens IT-infrastruktur og dermed sitte igjen med en løsning som er effektiv og produserer ønskede resultater. Blant virksomhetens IT-infrastruktur bør man også fokusere på sikkerhet. Dette innebærer å evaluere virksomhetens sikkerhetsrutiner, samt opptreden til virksomhetens nettverk og system. Dermed vil man hindre at aktører med uautorisert tilgang tilfører skader hos virksomheten. Samtidig verner man sensitiv informasjon fra å bli delt med uønskede parter (Epci University, u.å.).

Selve valideringen av modellen er også en viktig del av i evalueringsfasen. For at modellen skal implementeres, er den nødt til å testes og valideres. Dermed er man avhengig av å ha tilgjengelige teknologier som sikrer validering modellene. Testingen og valideringen gjøres for å sikre at den gir presise resultater. Det foreligger diverse teknikker og teknologier for å validere en modell, derav å foreta tester på individuelle sett med data, som for eksempel kryssvalidering (Wirth & Hipp, 2000). Ved å gjennomføre slike teknikker ved bruk av tilgjengelige teknologier, vil man kunne identifisere eventuelle feil i modellen, samtidig som at man forsikrer seg at den vil yte tilstrekkelig dersom den blir implementert.

## Ledelse

Ledelsen er ansvarlig for å sette mål for evalueringsfasen, og sørge for at dette blir gjort grundig. Ledelsen bør også definere tydelige kriterier, da det skaper forutsetninger for en prosess hvor modellresultatene blir målt opp mot forretningsmålene på en systematisk måte. Samtidig vil man da sette tydelige standarder til alle innad i virksomheten. Evalueringsfasen er en ekstremt viktig fase for å oppklare mulige svakheter i hele prosessen, og sørge for at disse blir fikset. Det vil videre være hensiktsmessig for ledelsen å lage en plan bestående av fremtidige steg. Blant dette inngår å produsere en eller flere lister over mulige handlinger som bør utføres i fremtiden, basert på resultatene fra evalueringen av modellene (Wirth & Hipp, 2000). Som kjent har modellene ulike egenskaper, noe som gjør at man bør lage flere lister slik at man til enhver tid har mulighet til å være handlingsdyktig (Ayele, 2020). Generelle eksempler på steg man kan gjennomføre, kan være å sjekke modellens modenhet opp mot virksomhetens forretningsmål, eller å sjekke om man bør foreta seg justeringer basert på dataen man har til rådighet.



## Kultur

Det å ha en kultur sentrert rundt datakompetanse, som er kritisk til resultatene, vil som i forrige fase være viktig også her. De ansatte må vurdere resultatene av fasen opp mot satte kriterier, og sørge for at disse blir oppfylt. Ifølge informant #1, skal det ofte *«ikke mer enn én gang til at dataanalysen feiler eller gir upålitelige resultater, før ledelsen stiller seg kritisk til å benytte analyse i beslutninger lenger»*. En kritisk kultur kan sørge for at alle ledd i prosessen samt resultatene blir evaluert nøye, noe som dermed minsker sannsynligheten for at resultatene blir upålitelige. Virksomheten burde forsøke å opparbeide en kultur der ansatte ikke bare er kritiske til egne resultater, men også andres. Organisasjonskulturer der ansatte generelt er kritiske til egne og andres resultater, men også villige til å lære bort kunnskap og ikke minst tilegne seg ny kunnskap selv, har en sammenheng med bedre resultater (Liu et al., 2021). I data-drevne virksomheter er det viktig med en kultur der ansatte gis rom til å eksperimentere og vurdere eget arbeid, og dermed lære av eventuelle feil (Anderson, 2015. s. 212).

## Strategi

Ledelsen i virksomheten er nødt til å ta en avgjørelse på hvordan de skal gå videre med prosjektet. Denne avgjørelsen gjøres basert på et samspill mellom resultatene fra evalueringen, og en analyse av mulige handlinger (Chapman, 1999). Dersom modellen anses som tilstrekkelig for å bli implementert i virksomheten, kan man begynne å planlegge en strategi for implementeringen av resultatet (Ayele, 2020). Hvis ikke, er virksomheten nødt til å foreta seg justeringer av modellen, eller analysere ytterligere data før de kan fortsette.

## Smidighet

Ved å kontinuerlig evaluere stegene som er tatt, vil man ha muligheten til å gå tilbake for å foreta seg endringer og justeringer. Dette vil være svært relevant dersom de foregående evalueringene markerer tydelige behov hos virksomheten (Chapman, 1999). Man vil ofte kunne oppdage problemer i denne fasen, som har ført til at prosjektet har resultert i modeller med lav/ingen verdi for virksomheten. Problemer her kan ofte spores helt tilbake til første fase av prosjektet, nemlig forretningsforståelse (Wirth & Hipp, 2000). Mål kan ha vært uklare, eller prosjektet kan ha startet før kritiske forretningsspørsmål var avklart. Problemer må avklares, og nødvendige tiltak må iverksettes for at de endelige resultatene skal tilføre virksomheten maksimal verdi. Ved å ha en fleksibel virksomhet, bestående av fleksible ressurser og systemer, vil evalueringene kunne bidra til å belyse hvilke tiltak som bør gjøres for å få maks effekt ut av implementeringen av de data-drevne makroanalysene. En fleksibel grunnstruktur hos virksomheten gjør det derfor tilnærmet problemfritt å gjennomføre en oppskalering av de aktuelle ressursene og systemene, dersom det skulle være nødvendig.

## 5.6 Implementering

Den avsluttende fasen i CRISP-DM, implementering, er en kritisk fase der modellen som er utviklet, implementeres i produksjon. Validering av modellen gjennom teknologi, skalering av løsningen, opplæring av de ansatte, implementering av modellen, overvåking og vedlikehold, endringsledelse, måling av ytelse og tillit til data er alle viktige aspekter som må tas i betraktning for å oppnå en vellykket implementering (J. S. Saltz, 2021).

## Data

Tillit til data er en essensiell faktor ved implementering. Man vil ikke være i god stand til å gjennomføre data-drevne makroanalyser, dersom man ikke har data med tilstrekkelig kvalitet. Basert på de foregående fasene skal man her ha endt opp med et pålitelig resultat. Dette skal være innsikt som direkte kan benyttes til å fremme endringer i virksomheten, eller alternativt modeller som skal integreres med f.eks. virksomhetens databaser for automatisk analyse av innkommende data (IBM, 2021g). Dersom det er førstnevnte, må innsikten ofte organiseres og presenteres på en måte som gjør at prosjektets kunde kan forstå seg på og benytte seg av den. Dersom det er sistnevnte, kan dette være mer komplekst, og kreve integrasjon av modellen med andre systemer i virksomheten (Wirth & Hipp, 2000). Dette vil kreve en plan for nøyaktig hvordan det skal gjennomføres. For disse modellene vil det også være viktig å planlegge hvordan de skal overvåkes og vedlikeholdes (IBM, 2021g).

## Metrikker

For modeller som skal integreres med andre systemer, vil det være viktig å forsikre seg om at de fremdeles produserer gunstige resultater i henhold til satte kriterier for modellen. Her burde man avgjøre hvilke faktorer som indikerer at modellen ikke presterer godt nok lenger (IBM, 2021g). Metrikker kan utnyttes for å måle dette. Avhengig av modellens formål, eksisterer det flere forskjellige typer metrikker som kan brukes, eksempelvis etter nøyaktighet, presisjon, og spesifisitet. For å velge de riktige metrikkene, må virksomheten først forstå hvilke utfordringer modellen skal løse, og hva som er de viktigste resultatene å måle. Videre må man sørge for at metrikker er realistiske og målbare, og at de kan gi innsikt i modellens effektivitet og ytelse. For modeller som skal gi direkte innsikt, kan det være mulig å benytte metrikker for å måle effektene bruk av modellen(e) har på virksomheten. F.eks. ble innsikten tatt i bruk i beslutningsprosesser? Hadde det positive/negative konsekvenser?

## Ferdigheter

Implementeringsfasen krever spesifikke ferdigheter for å sikre at modellen fungerer som den skal og gir de ønskede resultatene. I dette tilfellet, innebærer data-drevne makroanalyser ferdigheter innen teknologi, dataanalyse, programmering, og prosjektledelse. Derfor vil det være essensielt å ha eksperter med de riktige ferdighetene som kan bidra til å implementere modellen effektivt. Samtidig vil det være avgjørende å ha ansatte som er i stand til å tolke dataene, samt forstå resultatene som produseres av modellen.

## Teknologi

Dersom den utviklede modellen integreres med virksomhetens systemer, må virksomheten inneha teknologi for å foreta dette, samt for å overvåke modellen. Dersom prosjektet har ført til innsikt som skal benyttes direkte, må dette ofte visualiseres på en forståelig måte. Virksomheten må da inneha riktige datavisualiseringsverktøy for deres bruksformål. Det finnes mange ulike datavisualiseringsverktøy, fra mer generelle som Tableau til mindre verktøy som dekker nisjebehov (Lowe & Matthee, 2020). Ulike verktøy har ulike fordeler og ulemper, og hvilke(t) som skal benyttes må vurderes opp mot virksomhetens bruksbehov, samt virksomhetens evne til å benytte det.

## Ledelse

Dersom en modell skal integreres i virksomhetens systemer, vil ledelsen være ansvarlig for at en plan for hvordan dette skal gjøres blir etablert. Etter at modellen er satt i drift, er den nødt til å overvåkes og vedlikeholdes. Ledelsen er ansvarlig for å sikre at også dette skjer, samt for å sikre at kriterier til dens ytelse blir definert. Endringsledelse er en annen viktig faktor i implementeringsfasen av CRISP-DM, som ledelsen er nødt til å forholde seg til. Implementering av løsninger krever ofte en endring i virksomhetskulturen, noe som kan lede til motstand mot endringer blant ansatte og interessenter, og det er derfor viktig å få støtte fra de som er rammet, samt sikre at de ansatte er involvert i implementeringsprosessen (Thorsvik & Jacobsen, 2019, s. 391–392). Det er liten tvil om at en overgang til data-drevne makroanalyser kan framstå som ganske radikal for tradisjonelle virksomheter, og ved bruk av en rasjonell endringsledelse, kan dette bidra til at denne endringen blir mykere. Endringen kan framstå som mykere dersom man tydelig kommuniserer behovet for endringen, de potensielle fordelene ved den nye løsningen, samt involverer de ansatte i planleggingen (Todnem, 2005). Som et resultat av en suksessfull endring, vil man oppnå effektiviserte og forbedrede resultater. For å få en tilnærmet smertefri endring, bør man fatte en kort analyse der man gjør seg opp potensielle negative konsekvenser som enten kan oppstå under selve implementeringen, eller som en følge av den. Ved å kartlegge konsekvensene, vil man kunne handle raskt dersom det oppstår en uforutsett hendelse. Dette muliggjøres fordi man allerede har utviklet en plan. Som et resultat av denne strukturelle tilnærmingen, kan man oppleve en økt tillit og engasjement, samtidig som at sannsynligheten for å lykkes med implementeringen øker (Thorsvik & Jacobsen, 2019, s. 372–378).

## Kultur

I planen for implementering av en modell, burde ledelsen fokusere på opplæring blant de ansatte (Wirth & Hipp, 2000). Gitt at data-drevne makroanalyser er et relativt nytt fagfelt, er det rimelig å anta at svært få har kompetanse på dette området. Ved å ha tydelige rutiner for opplæring av de ansatte samt en læringsvillig kultur, bidrar man til å forsikre seg at alle involverte er i stand til å anvende modellen effektivt, og at det er tydelige retningslinjer for bruk av modellen. Dermed vil alle de ansatte vite hvordan man skal forholde seg til modellen.

Anderson (2015) forklarer at kulturen i virksomheter som ønsker å ta data-drevne beslutninger, også burde se verdien av god data, og ikke bare basere seg på intuisjon. Virksomheten må korrigere/avskaffe ledere som tar beslutninger som viser seg å være lite gunstige, til tross for at de i forkant ble presentert innsikt fra data som viste at dette mest sannsynlig ville skje. Ikke bare kan slike ledere direkte føre til dårlige beslutninger, men de kan hindre en åpen, samarbeidsorientert og læringsvillig kultur, hvor de beste, objektive og faktabaserte innsiktene vinner frem. Det vil være lite motiverende å jobbe som dataanalytiker, i en virksomhet som ikke benytter seg av arbeidet ditt (Anderson, 2015, s. 214–215).

## Strategi

Strategien i implementeringsfasen bør fokusere på en kombinasjon av datavitenskapelig ekspertise og ekspertise på fagområdet, i dette tilfellet data-drevne makroanalyser. Ofte sliter virksomheter med å behandle ustrukturerte data, noe som gjør det svært

hensiktsmessig å ha ansatte som har en solid kunnskapsbase innenfor datavitenskap og data-drevne makroanalyser. Dersom man ikke besitter ansatte med relevant ekspertise innenfor de to fagfeltene, vil man oppleve skjeve resultater, samtidig som at man ikke får en reel verdi fra resultatene man sitter igjen med (Kang et al., 2022). Derfor er en kombinasjon av ekspertise meget sentralt, slik at man ikke feilprioriterer ressursene man har tilgjengelig. Dette krever en god planlegging, og det er derfor viktig å ha en tydelig strategi for hvordan du ønsker å rasjonere de tilgjengelige ressursene. I tillegg bør strategien her utformes basert på forretningsforståelse, for å sikre implementering som samsvarer med virksomhetens behov.

## Smidighet

Dersom integrerte modeller skulle slutte å prestere etter etablerte kriterier, lønner det seg med en forhåndsbestemt plan over hva som da burde gjøres. I noen tilfeller er det nok med små endringer i modellen, mens det i andre tilfeller krever at man må starte et på bunn av og bygge en helt ny modell (IBM, 2021g). En etablert plan for hva som bør gjøres når dette skjer, sikrer en smidig håndtering. Overvåking og vedlikehold av modellen er essensielt for å sikre at den gir tilstrekkelige resultater over tid (Chapman, 1999). Dermed er man også i stand til å effektivt foreta justeringer og tilpasninger i modellen (Wirth & Hipp, 2000). Justeringer og tilpasninger som typisk oppstår, innebærer blant annet å tilpasse seg endringer i dataene som modellen bruker. Det kan også være aktuelt for ledelsen å foreta endringer i programvaren som støtter modellen, dersom det oppstår endringer i produksjonsmiljøet.

## 6 Diskusjon

I dette kapitlet vil vi se tilbake på prosessen som helhet og på funn og resultater som vi kom frem til. Videre skal undersøke disse opp mot teori, og informantenes utsagn for å besvare problemstillingen. Vi kommer til å drøfte rundt valg av forskningstilnærming og bruk av kilder, før vi videre går over på resultat og funn.

### 6.1 Forskningstilnærming

Innledningsvis i prosessen ble vi forespeilet av oppgavestiller en veldig åpen og lite konkret oppgave, men med et tydelig problemtema. Dette satte mye føringen for hvordan vi valgte å initiere forskningstilnærmingen vår. DSRM ble raskt valgt etter at vi konkluderte problemstillingen, basert på at vi ønsket å utvikle et rammeverk. Dette krevde en prosess som tok for seg informasjonsinnhenting og forståelse, i tillegg til en praktisk del som innbefattet demonstrasjon og tilbakeføring til tidligere steg, for å utføre prosessen på nytt frem til et tilfredsstillende resultat var oppnådd. Denne formen for forskning ble i våre øyne ansett som hensiktsmessig da den la vekt på at vi først ville få en dypere forståelse av tematikken og virksomhetsdriften, før vi satte i gang med prosessen å utforme kravene til rammeverket og etter hvert rammeverket i seg selv.

Selv om vi raskt konkluderte med valget av DSRM, undersøkte vi hvordan en mer tradisjonell kvalitativ, empirisk studie ville fungere. Dersom vi hadde benyttet oss av en slik form for forskningstilnærming, ville denne iterative delen av forskningen forsvunnet og vi hadde måttet lent oss mer på å treffe første gangen vi formulerte kravene. Dette betydde at en mye mer grundig studie måtte bli gjennomført og at vi i større grad hadde gjennomført prosessen uten å få noen form for bekreftelse underveis. Oppgaven hadde nok også hatt en større forankring i teori og mindre praktisk syn på gjennomføringen av rammeverket. At besvarelsen hadde hatt et større preg av teori og mindre praksis ville muligens ikke påvirket den negativt. Vi anså det totalt sett som bedre for modellens utforming at den hadde et praktisk syn på gjennomføringen.

### 6.2 Valg og bruk av kilder

Siden vi hadde valgt en forskningstilnærming som vektla en dypere forståelse av tematikken, var valg og bruken av kilder viktig. Det har gjennomgående i oppgaven vært viktig å ha en kritisk holdning til kildene som har blitt tatt i bruk. Til tross for at majoriteten av kildene kan fremstå som legitime, eksisterer det fortsatt en stor variasjon i kvaliteten på artiklene, samt forskernes omdømme. Siden vi ønsket å skrive en oppgave som forankres i faglig gode kilder, ville vi tjene stort på å foreta en evaluering av diverse faktorer. Eksempler på aktuelle faktorer som vi evaluerte før bruk, var forfatterens kompetanse, forlagets eller institusjonens omdømme, artikkelens alder, bruk av metode, og forskningstilnærming. Dermed forsikret vi oss at besvarelsen vår bygde på sikre, objektive og relevante kilder. I tillegg til å foreta seg disse evalueringene hadde vi en konsekvent nøyaktig holdning til kildehenvisning, gjennom å gjennomgående korrekt sitere og referere til kildene som vi anvendte. Ved å gjøre dette vil eksterne personer kunne se gjennom kildene våre, og deretter foreta seg evalueringer av kildens

nøyaktighet og legitimitet, og samt ha mulighet til å etterse at informasjonen vår er riktig. Bruken av fagfelleverderte kilder er en grunnpilar i vitenskapelig forskning og akademisk skriving. Dette gjør at vi ønsker å poengtere betydningen av å bruke slike kilder, samt viktigheten av å være kildekritisk. Primærkilder som vi brukte i besvarelsen vår, var hovedsakelig vitenskapelige tidsskrifter, bøker som er utgitt av anerkjente forlag, og empiriske studier. Vi benyttet oss av andre typer kilder utover dette, men disse ble grundig vurdert. Sekundærkilder besto av intervjuer av fagpersoner. Bruken av disse kildene gjorde at vi forankret arbeidet i legitimitet, samt høy kvalitet. Dette siden kildene bygger på vitenskapelig metode og forskningspraksis.

### 6.2.1 Primærkilder

Primærkilder symboliserer allerede stadfestet vitenskapelig arbeid, noe som gjorde tidligere vitenskapsartikler til stor hjelp for arbeidet vårt. Ved bruk av akademiske og fagfelleverderte kilder, bygget artikkelen vår på etablert kunnskap, samt tidligere forskning gjort av andre. Dette muliggjorde at vi kunne henviser til tidligere fenomener som var oppdaget, trekke paralleller, utfordre eller bekrefte tidligere funn, identifisere eventuelle mangler. Tidligere forskningsartikler baseres også på metoder, resultater og konklusjoner fra tidligere studier, noe som gjorde at vi var i stand til å sammenligne og evaluere egne funn opp mot tidligere funn. Denne sammenlikningen gjorde også at våre egne funn ble mer reflektert. Som et resultat av å forankre besvarelsen vår i tidligere fagfelleverderte kilder, ble kvaliteten på arbeidet generelt bedre, samtidig som det er forsvarlig å anta at leserens tillit til oss som forfattere styrkes.

Generelle utfordringer vi møtte på ved bruken av primærkilder, omhandlet tilgjengelighet. Tidvis kunne det være vanskelig å finne relevante primærkilder om problemstillingen vår. Dersom vi fant artikler som var relevante, kunne det oppstå problemer med konteksten. Med dette mener vi at deler av informasjonen i artikkelen var relevant, og dermed kunne brukes, men at det samtidig var vanskelig å finne denne informasjonen. Til tross for at vitenskapelige artikler gjerne gjennomgår en grundig vurderingsprosess der aktuelle eksperter innenfor fagfeltet foretar seg grundige evalueringer av innholdet, kan artiklene bære preg av å være upresise eller subjektive. Enhver person har gjerne en særegen tolkning av diverse fenomener man kommer over, noe som kan påvirke artiklene. Dersom vi oppdaget at forfatteren ikke var en anerkjent fagperson innenfor sitt respektive fagområde, foretok vi oss derfor eksterne undersøkelser i ytterligere litteratur. Ved å gjøre undersøkelser i ytterligere litteratur forsikret vi oss forsvarlig informasjon som var relevant for besvarelsen vår.

En annen utfordring vi støtte på var hvor gammel kilden var og eventuelt om den var utdatert. Mye innen teknologi har endret seg drastisk de siste årene og kanskje spesielt under og etter covid-pandemien. Vi drøftet derfor grundig rundt dette og kryssjekket gjerne kilder før vi valgte å ta de i bruk, slik at vi kunne forsikre oss at informasjonen ikke var utdatert eller misvisende. Dette spesielt rundt kravutvikling og senere skissering av rammeverket innenfor de områdene som krevde teknologi og kunnskap. Områdene som omhandlet organisasjon og endringsledelse bygger i stor grad på mer eller mindre standardiserte prinsipper, og vi kunne dermed tillate bruken litt eldre kilder her.

### 6.2.2 Sekundærkilder

Vi brukte intervjuer som sekundærkilde i besvarelsen vår. Ved å bruke intervjuer som kilde, kunne vi tilegne oss en dypere forståelse av aktuelle temaer som vi alt diskuterte. Enkelte emner innenfor problemstillingens begrensninger var mangelfulle, noe som gjorde det svært nyttig og verdifullt å tilegne oss ytterligere informasjon basert på fagpersoners erfaringer og kompetanse. Intervjuer gjorde det også mulig å gå i dybden på områder vi ellers hadde lite kunnskap om, samt å skaffe oss skreddersydde svar på lite omtalte fagområder. Dermed var vi i stand til å tilføye nye dimensjoner til besvarelsen vår, da vi kunne sammenligne allerede anvendt informasjon, og data, opp mot sitater fra foretatte intervjuer. Dette gjorde også at vi var i stand til å trekke paralleller, samt utfordre eller bekrefte tidligere funn. Ved å anvende intervjuer på en slik måte, var det hensiktsmessig å kategorisere de som sekundærkilder, da sekundærkilder gjerne skal fungere som tilleggsinformasjon for primærkilder. Likevel er det enkelte utfordringer knytte til bruk av intervjuer som kilde. Informasjon og utsagn fra informanter kan enten være subjektive, eller bære preg av subjektivitet, da de baserer seg på vedkommendes erfaring og bakgrunn. Derfor må informantens bakgrunn tas i betraktning når man evaluerer aktualiteten til informasjonen slik at man minimerer risiko for å bli feilinformert, eller å få irrelevant informasjon.

### 6.3 Drøfting av resultat

Vi valgte tidlig at vi skulle utvikle et rammeverk. Dette i stor grad forankret i at vi gjennom litteratursøket konkluderte med at der var lite teori som tok for seg kombinasjonen tekniske og organisatoriske krav, samtidig som at vi sto ovenfor et problem formulert av oppgavestiller. Selve grunnen til at vi valgte å gå for å utvikle et rammeverk kontra en metode baserte seg på rammeverkets mulighet til å være mer fleksibelt. Metode derimot er bygget på prinsipper eller regler for fremgangsmåte. Basert på at vi skulle ta for oss både tekniske og organisatoriske faktorer som krever en del tilpasninger ut ifra virksomheten, anså vi det derfor som mer hensiktsmessig å gå for en mindre standardisert fremgangsmåte. Dette valget ville ha en stor påvirkning på hvordan sluttresultatet vårt ville bli. Istedenfor å skulle ende opp med en standardisert fremgangsmåte, ville vi ende opp med oversikt eller anbefaling til hva som burde ligge til grunn.

Rammeverkets oppbygning ble fastslått basert på hva vi ønsket at rammeverket skulle ta for seg. Med grunnlag i problemtematikken kunne vi konkludere med at virksomheter ofte forsøkte å gjennomføre endringer uten at det var forankret i et reelt forretningsbehov. Et gjennomgående tema i samtalene våre med ansatte i twoday Avento var at kundene deres ønsket hjelp til å digitaliseres eller bli mer data-drevne, uten å nøyaktig ha klart for seg hva de ønsket, hvorfor de ønsket det, og hvordan de så for seg å gjennomføre det. Uten å ha dette klart for seg, er det vanskelig å utvikle strategier for hvordan det skal gjennomføres. Dette ga hovedsakelig grunnlaget for at rammeverket tar for seg tre hovedmomenter: Virksomhetens utgangspunkt og deres motivasjon for å bli data-drevet, prosessen rundt det å skulle bli det, og hva som kreves av virksomheten for å gjennomføre. Det vi også raskt merket manglet i litteraturen var viktigheten ved gevinstrealisering i virksomheter. Med bakgrunn i det kombinerte teknologiske og organisatoriske perspektivet vi har tilegnet oss gjennom studiet vårt, var det slående hvordan ideene om verdirealisering enten ikke ble anerkjent, eller slått sammen til et lite steg uten en forklaring rundt hvordan det burde gjennomføres. Dette er ikke noe vi bare har adressert, men vi har også tilbudt en måte å tenke rundt det. Vi

så behovet for å knytte gevinstrealisering med teknologien, og har følgende gjort dette gjennom det ferdige resultatet.

Siden gevinstrealisering var sentralt for oss som digitale forretningsutviklere, anså vi det også som naturlig å innbefatte dette i rammeverket vårt. Dette i kombinasjon med implementering av teknologi, altså prosjektledelse. I prosessen med å velge hva vi skulle basere rammeverket på var vi også innom å ha et kompetansebasert rammeverk. Denne typen rammeverk hadde som utgangspunkt å gjøre en virksomhet mer kompetent og tilpasningsdyktig. Dette er selvfølgelig element som er viktig i forhold til beslutningstaking. Gruppen valgte likevel å ikke ta dette med i utformingen av rammeverket, da vi følte vi fikk inn elementer av tilpasningsdyktighet og kompetanse på en god og dekkende måte gjennom de åtte virksomhetsdimensjonene presentert av Buitelaar (2018).

Da kriteriene for prosjektledelse og gevinstrealisering var satt, la dette også føringen for hvordan vi videre skulle gjennomføre prosesser i rammeverket. Ved undersøkelse av modeller for Data Mining kom som tidligere beskrevet KDD og SEMMA opp i tillegg til CRISP-DM. KDD og SEMMA hadde en fremgangsmåte rettet veldig opp mot selve uthenting av data, og tok i liten grad for seg hvordan dette skulle brukes opp mot virksomheten. Dette gjorde derimot CRISP-DM gjennom den innledende fasen forretningsforståelse. Dette var grunnen til at vi valgte å gå for CRISP-DM over de andre metodene. Vi kunne nok lagt til dimensjoner om forretningsforståelse til KDD og SEMMA, men dette ble ansett som unødvendig da CRISP-DM alt hadde dette innbefattet i modellen.

Likevel anså vi CRISP-DM som en mangelfull modell i forhold til om virksomheter evnet å skulle bli data-drevet og gjennomføre makroanalyser. Med bakgrunn i dette ville vi derfor videreutvikle CRISP-DM til å problematisere dette for virksomhetene, og valgte derfor å legge til ytterligere analyser i fasen forretningsforståelse. Målet med dette var å tvinge virksomheter til å bedre forstå hva deres reelle behov er, hvordan de kan løse det, og hva eventuelle gevinster vil bli. Ønsket var at virksomheter bedre skal være i stand til å tenke på langsiktige konsekvenser, og ikke ta beslutninger basert på et følt, men nødvendigvis ikke reelt behov. Endringene som skulle implementeres må også ha større nytte enn kostnad. Kombinasjonen vår av fasene i CRISP-DM med de åtte virksomhetsdimensjonene, synliggjør omfanget av et slikt prosjekt. Gjennom dette dekker rammeverket samtlige viktige faktorer som må vurderes og aktiviteter som må gjennomføres, langs hele prosessen av et slikt prosjekt. Bruken av en generell modell som CRISP-DM, samt åtte identifiserte felles virksomhetsdimensjoner, gjør at alle virksomheter tenkelig kan ta i bruk rammeverket. Basert på dette, vil de få et mer realistisk bilde av det faktiske omfanget, og hvordan det vil påvirke deres virksomhet. Rammeverket vil fungere som et nyttig verktøy i planleggingsfasen av slike prosjekter, enten de har utviklet data-drevne makroanalyser før, eller dersom de forsøker å utvikle det for første gang.

Rammeverket er generelt utformet, da vi ønsker at nær alle virksomheter som ønsker å utvikle data-drevne makroanalyser skal kunne ta det i bruk. Det er ikke ment å gi en fasit, men fungere som et forslag for hvordan virksomheter kan gjennomføre det. Tanken er at alle skal kunne benytte det til å forstå omfanget, samt planlegge for det. De overordnede aktivitetene presentert i CRISP-DM må gjennomføres i en eller annen form for å utvikle data-drevne makroanalyser. Vi har gjennom arbeidet forsøkt å besvare hvordan dette kan gjøres, og hva som må vurderes i perspektiv av et makroanalytisk



prosjekt. Totalt sett ønsker vi at rammeverket skal sørge for at flere virksomheter suksessfullt evner å gjennomføre slike prosjekter, at det vedtatte prosjektet har en reell verdi, og til slutt at virksomheten faktisk klarer å realisere denne verdien.

Det er selvfølgelig noen utfordringer knyttet til data-drevne makroanalyser, dette er spesielt når det kommer til data og input. Disse fasene er innledende avhengig av at mennesker overser og utfører prosesser. Til tross for at data-drevne makroanalyser kan være meget nøyaktige, er det likevel rom for at det kan oppstå feil, misforståelser eller unøyaktigheter. Eksempler som kan lede til dette kan være at dataene som anvendes ikke er representative, eller at det mangler vesentlige opplysninger som kan påvirke analysen. Dermed er det viktig å være bevisst på begrensningene ved dataene man besitter, og å sikre god datakvalitet. Data-drevne makroanalyser kan også kreve mye tid og ressurser for å utvikle, samt analysere store mengder data. Dette gjør at det kan være kostbart å skaffe og vedlikeholde nødvendig teknologi og programvare for å utføre de aktuelle analysene. Ved bruk av store datamengder kan man oppleve bekymringer knyttet til personvern og datasikkerhet. Derfor er det viktig å være oppmerksom på disse risikoene, slik at man kan gjennomføre daglig drift på en forsvarlig måte. Kvaliteten ved dataene som anvendes i makroanalysene kan framstå som en utfordring. Dataene kan enten være mangelfulle, ha feil eller være utdaterte, noe som kan resultere i unøyaktige, eller feilaktige konklusjoner og beslutninger. Datakvalitet kan være svært krevende å sikre, da det avhenger av brukstilfellet for dataen. Sikring av datakvalitet kan også innebære innsats fra mange ulike ansatte, med varierende grader av datakompetanse. Bruk av data-drevne makroanalyser krever svært avansert teknologi og kompetanse, noe som kan gjøre det utfordrende å holde seg oppdatert på nye teknologier, metoder og løsninger. Stadige endringer kan være en utfordring for data-drevne makroanalyser. Teknologier utvikler seg i et skyhøyt tempo, noe som kan gjøre det vanskelig å holde følge. Samtidig kan en ustabil situasjon i verden gjøre at uforutsette hendelser oppstår, med store ringvirkninger. Det er mange ulike faktorer som kan spille inn, noe som kan gjøre at det er vanskelig å forutse om data-drevne makroanalyser er like aktuelt i framtiden, eller om det er kommet en enda mer kompleks løsning.

Likevel dersom data-drevne makroanalyser utvikles og implementeres på riktig vis vil det bidra til mange positive muligheter og styrker til en virksomhet. Ved bruken av data-drevne makroanalyser kan man bedre tilpasse seg endringer blant omgivelsene. Denne tilpasningen skjer gjennom å identifisere trender og mønstre i markedet, eller globale omgivelser. Dermed vil virksomheter være bedre rustet til å håndtere potensielle utfordringer. Data-drevne makroanalyser kan også bidra til innovasjon og nyskaping. Ved å analysere og håndtere innsamlet data, kan man identifisere potensielle muligheter, gjennom behov eller mangler blant konkurrenter og markedet. Dette gjør at man kan utvikle nye løsninger. Data-drevne makroanalyser gir en presis og objektiv tilnærming av analysene som blir foretatt av omgivelsene, sammenliknet tradisjonelle analyser. Dette skyldes at dataene som samles inn, og videre analyseres kun er basert på håndfaste tall og fakta, og det er dermed mindre rom for personlig tolkning og resonering. Data-drevne makroanalyser kan håndtere store mengder data på en mer effektiv måte, sammenliknet tradisjonelle analyser. Som et resultat muliggjør dette å analysere en større mengde informasjon på kortere tid, samt å synliggjøre forskjellige mønstre og trender, som tradisjonelle analyser derimot ville hatt vanskelig for å avsløre. Ved å avsløre slike trender og mønstre, vil dette gi virksomheten et større konkurransefortrinn. Dette skjer på grunn av deres evne til å forutse nye trender før andre konkurrenter, og dermed reagere proaktivt og tilpasse seg situasjonen tidligere. I motsetning til tidligere analyser, kan data-drevne makroanalyser være tidssparende for virksomhetene som anvender

teknologien. Man vil oppnå en effektivisering, gjennom at man automatiserer dataanskaffelsen, samt selve analyseprosessen. For virksomheten betyr dette at de oppnår en reduksjon i både tid og ressurser, og kan dermed prioritere andre arbeidsoppgaver eller behov.

### 6.3.1 Drøfting av utvikling av kravene til rammeverket

Ved utviklingen av stegene i kravhåndteringen, ønsket vi å forankre disse i terminologien om MDRE. Fordi det ikke er et universelt svar på hvordan virksomheter spesifikt skal foreta seg utvikling av data-drevne makroanalyser, gjorde MDRE det mulig å liste opp generelle forutsetninger som må ligge til grunn for å lykkes. Dermed ville vi sitte igjen med en fremgangsmåte som foretar seg utvikling av løsninger rettet mot større markeder, og dermed ikke spesifikke interessenter. Basert på denne fremgangsmåten var vi i stand til å benytte fire steg i kravhåndteringen, bestående av kravinnhenting, kravanalyse, kravdokumentasjon, samt krav-verifisering og validering. De utviklede kravene er selve fundamentet i løsningen vår, og dersom vi ikke hadde fulgt en prosedyre for å utvikle disse, ville vi muligens ikke forstått omfanget av alt som måtte inkluderes. Dersom kravene våre ikke hadde gjennomgått stegene kravhåndtering består av, ville de ikke blitt forankret i noen form for legitimitet. Med legitimitet mener vi at de utformete kravene er spesifikke, og dekker de reelle behovene tilhørende rammeverket, samt dens interessenter. Med utgangspunkt i kravhåndteringen prøvde vi å utvikle krav som samsvarte med prosessmodellen CRISP-DM og Buitelaars åtte dimensjoner. Gjennom denne kombinasjonen sikret vi oss en utforming av krav som både tok for seg organisatoriske og tekniske faktorer. Til tross for dette, oppsto det en tidvis skjev vektning i utdypning av kravene. På grunn av at vi innledningsvis ønsket å ta utgangspunkt i den nevnte kombinasjonen, var det utfordrende å få de allerede etablerte kravene til å samsvare med CRISP-DMs faser og Buitelaars åtte dimensjoner. Som et resultat, ledet dette til at vi måtte evaluere kravene, og foreta oss justeringer deretter. Dermed ble grunnlaget til enkelte av kravene i rammeverket svekket, til tross for at andre fremstår som veldefinerte og solide, noe som leder til den skjeve fordelingen.

### 6.3.2 Drøfting av funnene i rammeverket

Videre skal vi drøfte funnene vi kom frem til gjennom skissen til rammeverket. Dette gjøres etter CRISP-DMs inndeling av faser.

#### Forretningsforståelse

Fasen forretningsforståelse gjennomgikk en større omstrukturering mellom utviklingen av krav og skisse av rammeverket. Bakgrunnen for dette er beskrevet i kapittel 3.5 og 3.6. Vi valgte å inkludere analysemetodene PESTEL, SWOT og BDN i forretningsforståelsesfasen fordi de dannet et solid grunnlag for forretningsforståelse. PESTEL ble lagt til fordi det er et enkelt rammeverk, som har en tydelig struktur og er enkelt å følge. Det kan gjøre det mulig for en virksomhet å oppdage forretningsmuligheter og utnytte dem fullt ut. Disse fordelene gjelder dersom analysen gjennomføres riktig. Likevel er det noen ulemper ved PESTEL-analysen. Dersom brukere forenkler mengden data som brukes til beslutninger, kan det her være lett å bruke utilstrekkelig data. I tillegg, for å være effektiv, må PESTEL gjentas med jevne mellomrom, og dette kan være noe virksomheter ikke ønsker å bruke ressurser på. SWOT ble naturlig å ta med som en forlengelse av PESTEL.

Basert på undersøkelser vi har gjort, kombinert med funn vi har oppdaget, spiller digital modenhet en sentral rolle for å lykkes med utviklingen av data-drevne makroanalyser. Det vil si at manglende digital modenhet hos virksomheter gjør det tilnærmet umulig å være i stand til å drive et makroanalytisk prosjekt. Ved forståelsen vi har tilegnet oss gjennom dette prosjektet, er det særlig fire utfordringer som utgjør typiske mangler blant digital modenhet hos virksomheter. Disse utfordringene er mangel på eierskap til data, og mangel på kompetanse og digital modenhet. Dette i tillegg til mangel på realistisk tidsperspektiv og langtidsperspektiv, samt en ambisiøs, men lite realistisk tilnærming til digitaliseringsprosjekter.

Med mangel på eierskap til data menes at virksomheten enten ikke har tilgang på data, vet hvordan de generer data, eller hvordan de forvalter data. Eierskap til data er essensielt for å lykkes med data-drevne makroanalyser, da dataen fungerer som byggeklosser i prosjektet. Derfor er det aktuelt å understreke viktigheten av at virksomheter gjennomgår faser som forretningsforståelse, dataforståelse og dataforberedelse som en del av et såpass omfattende prosjekt. Disse fasene tvinger de til å ta eierskap til data, kartlegge kompetanse, definere strategier og definere behov, noe som tilrettelegger en stor del av forutsetningene for å lykkes.

Mangel på kompetanse og digital modenhet er en annen typisk utfordring for virksomheter. For å lykkes med generelle prosjekter er man nødt til å besitte en kombinasjon av kompetanse og ressurser. Det vil være vanskelig å gjennomføre prosjekter dersom man kun besitter en av delene. Dersom man kun har kunnskap, vil man ikke ha noe å bygge på, og dersom man kun har ressurser, vil man ikke vite hvordan man skal gå fram. Derfor er det hensiktsmessig å forsikre seg at virksomheten besitter begge deler når man innledningsvis vedtar et slikt prosjekt. Ved å besitte kunnskap vil man ha evnen til å kunne oversiktlig planlegge steg, vite hva man trenger av ressurser, vite hvilke valg man bør ta, identifisere eventuelle behov, og hvordan man skal forholde seg til andre interessenter og aktører. Ved å evaluere den digitale modenheten tilhørende virksomheten, vil man finne ut om virksomheten er egnet til å gjennomføre et digitaliseringsprosjekt. Evalueringen av den digitale modenheten fokuserer på IT-ressursene som virksomheten råder over. Dermed vil man være i stand til å identifisere hvilke ressurser som anvendes, og hvordan disse anvendes. Ved å vite «hva», og «hvorfor», vil det være lettere å identifisere eventuelle behov for hva som videre trengs av virksomheten i digitaliseringsprosjektet. Det er derfor viktig at dette gjøres innledningsvis i prosessen som en del av de tre analysene, slik at man kartlegge hva bedriftens nåværende punkt er og hva de eventuelt trenger av kunnskap og erfaring for å være god nok rustet til å gjennomføre utviklingen og implementering.

En ambisiøs, men lite realistisk tilnærming til digitaliseringsprosjekt er en tydelig utfordring hos virksomheter. Denne utfordringen bærer gjerne preg av forrige punkt vi diskuterte, mangel på kompetanse og digital modenhet, samt neste punkt som skal diskuteres, realistisk tidsperspektiv. Å være ambisiøs ved inngangen av enhver type prosjekt er vanlig, og skyldes ofte mangel på kunnskap. Denne kunnskapen bærer preg av mangel på hva som kreves, og hva som trengs for å gjennomføre slike prosjekter. For å unngå dette, vil det være essensielt å ta seg god tid før prosjektet til å drøfte rundt hva som kreves i et slikt prosjekt. Typiske ting som bør drøftes er organisatoriske ressurser man har til rådighet, både i form av tid, økonomiske midler, ressurskapital, og bemanning. Videre bør man legge en plan for hele prosessen, der man spesielt vektlegger hvilke fremgangsmåter man ønsker å bruke, strategier, delmål og mål, samt resultater man ønsker å sitte igjen med etter endt prosjekt. Ved å disponere tilstrekkelig

tid til å planlegge gjennomførelsen av prosjektet, sikrer man seg en mer balansert og forutsigbar gjennomføring.

Mangel på realistisk tidsperspektiv er en utfordring som har blitt belyst gjennom prosjektet vårt. Denne utfordringen henger sammen med at virksomheter har en lite realistisk, men ambisiøs tilnærming til digitaliseringsprosjekter. Likevel er det verdt å utdype denne utfordringen, da tid er en dyrebar faktor i ethvert prosjekt. Ut ifra kunnskap tilegnet gjennom prosjektet vårt, kommer det fram at virksomheter ofte undervurderer tiden det tar å gjennomføre et digitaliseringsprosjekt. Dermed risikerer man å få et ufullstendig, eller i verste fall, totalt mislykket prosjekt. Mange virksomheter kaster seg ut i prosjekter uten å gjøre tilstrekkelig forarbeid. Hvis man kaster seg ut i et prosjekt uten et tydelig mål, visjon og plan, vil sannsynligheten for å lykkes synke drastisk. I mange tilfeller er det bedre å levere et prosjekt som er fullstendig, men over tidsfrist, enn å levere et mislykket prosjekt til tide. Videre har det også kommet fram en mangel på et langsiktig perspektiv. Med dette menes at virksomheter ønsker å høste resultater, gjerne før resultatene gir en reel verdi. Dette skyldes en kombinasjon av utålmodighet og ivrighet, som gjør at man blir blinde på verdien løsningen vil gi i framtiden.

## Dataforståelse

I fasen dataforståelse, er den overordnede aktiviteten i CRISP-DM basert på at dataen som skal benyttes til analyseformål skal samles inn. Kravene vi definerte til fasen utgjorde overordnet at vi måtte utvikle et forslag til hvordan virksomheter kan sikre innsamling av korrekt data, samt hva som må være på plass, eller hva som kan påvirke gjennomførelsen av denne fasen. I perspektiv av makroanalyser er dette en av de mest vikende fasene, sammenlignet med tradisjonelle dataanalyser, ettersom omfanget av data man kan benytte nærmest er ubegrenset. Spesielt gjeldende var dette for datadimensjonen i rammeverket. Eksempelvis kan det være værdata for å optimalisere forsyningskjeden, eller internasjonal data om f.eks. prisutvikling i land som kan påvirke virksomheten din.

Basert på litteraturen vi fant fremstod det ikke som at det eksisterte en konsensus rundt hvordan primært ekstern data skulle vurderes. Det var tydelig fra litteraturen og intervjuene at datakvalitet er essensielt. Basert på Sidi et al. (2012) etablerte vi tidlig i litteraturkapittelet at det finnes enormt mange dimensjoner av datakvalitet, og at hva «*god datakvalitet*» er avhenger av bruksformålet. Basert på et allerede etablert bruksformål fra fasen forretningsforståelse, konkluderte vi med at datakvalitet må vurderes etter dette. Videre foreslår vi at nye datakilder burde vurderes etter de 12 faktorene presentert av Anderson (2015. s. 51-52). Disse to vurderingene burde igjen kombineres i en strukturert tilnærming til vurdering, lik den presentert av Legner et al. (2022). Slik dannet vi en foreslått prosess for utvalg av data basert på eksisterende litteratur, som svarer på kravene vi definerte.

Basert på etablert forståelse understreket vi at datakvalitet kan måles gjennom metrikker, men at dimensjonene det måles etter må samsvare med tenkt brukstilfelle. Basert på intervju med informant #1, fant vi ut at det er viktig at ansatte som arbeider direkte med data, har tilstrekkelige ferdigheter for å evne å forstå hva dataen gir informasjon om. De ansatte må forstå hvilke data som kan brukes, og hva som evt. bør forkastes. Når det kommer til teknologien, fant vi ut at det eksisterer enormt mange verktøy, og at det er vanlig å benytte såkalte ETL-verktøy i prosessen med å hente ut,

transformere, og laste data inn i virksomhetens egne systemer. Kherdekar & Metkewar (2016) sier at det ikke eksisterer et universalt verktøy som dekker alle virksomheters behov. Vi eksemplifiserte noen mulige verktøy, men konkluderte med at det ville være for omfattende å ta for seg ethvert verktøy i detalj, både med tanke på lengden av oppgaven, men også kunnskapen vi besitter om temaet. Det viktigste er at valgt teknologi samsvarer med virksomhetens behov som ble etablert innledende fase i prosjektet, og at ulike virksomheter burde ta ulike faktorer i betraktning når de velger hvilke verktøy de skal benytte.

Ved ledelse baserte vi oss på funnene til Legner et al. (2022). Med utgangspunkt i hans funn, foreslår vi en liste over roller som burde utfylles i virksomheten. En helt ny rolle er dataspeider, som har hovedansvar for å identifisere, vurdere og forvalte ekstern data. Det var lite faglitteratur å finne rundt en slik rolle i praksis, men en lignende rolle ble blant annet foreslått av McKinsey (Aaser & McElhaney, 2021). Basert på vår kunnskap om virksomhetskultur, understøttet av Anderson (2015), fant vi det hensiktsmessig å foreslå en innovasjonsfremmende kultur. En innovasjonsfremmende kultur bidrar til eksperimentering og testing i et såpass lite utforsket område som data-drevne makroanalyser. I henhold til strategi, konkluderte vi med at virksomhetens strategi måtte være tydelig definert og kommunisert. Vi har tidligere diskutert at behovet for data avhenger av virksomhetens behov. En strategi de ansatte forstår, er dermed et grunnleggende behov her. Ved smidighet fant vi ut, basert på intervju med informant #4, at de største utfordringene vedrørende datakilder er overvåkning, og å møte eventuelle endringer. Dermed konkluderte vi med at gode prosesser for overvåkning og forvaltning er essensielt, spesielt versjonshåndtering som understrekes av Legner et al. (2022).

Oppsummert har denne fasen bidratt til å besvare problemstillingen gjennom å diskutere aktiviteter og faktorer som samlet sett oppfyller de aktuelle kravene. Dette omhandler hovedsakelig hva som må adresseres i de ulike virksomhetsdimensjonene for å sikre innsamling av korrekt data. Vi kunne muligens foretatt oss en ytterligere utdypende forklaring av potensielle datakilder, slik som de vi har forklart under innsamlingslaget av dataarkitektur i litteraturkapittelet. Likevel er dette noe som virksomheter burde vektlegge selv, da ulike kilder kan være relevante for ulike virksomheter. I Norge har vi f.eks. data.norge, som kun er relevant for norske virksomheter. Basert på litteraturen virker det å være et voksende felt, hvor stadig flere kildeleverandører og potensielle nye kilder dukker opp. Dette ville muligens ledet til at informasjonen vi anvendte, raskt ville blitt utdatert.

## Dataforberedelse

I dataforberedelsesfasen utgjorde kravene at vi måtte ta for oss hvordan data skulle forberedes for makroanalyser, samt hva dette ville kreve av virksomheten. Basert på de foretatte intervjuene, og Chu et al. (2016), er dette en kritisk, samt tidvis vanskelig oppgave. Vi baserte oss på vår etablerte forståelse av de mange ulike dimensjonene av datakvalitet, og at disse utgjør ulike fokusområder virksomheter burde prøve å forberede. I datadimensjonen tok vi for oss selve prosessen rundt forberedelse av data. Vi baserte oss på Chu et al. (2016) sine tre spørsmål for «hva, hvordan og hvor» feil skulle både oppdages og repareres. Dette kan være en meget teknisk prosess, som trolig krever mer spesialisert kunnskap om temaet, enn hva vår egen bakgrunn tilsier at vi er i stand til å forstå. Dette skapte noen begrensninger for oss, men vår anvendte kilde dekket det mest grunnleggende. Vi beriket også underpunktene med innsikt fra andre

kilder. Basert på den foregående tilbakemeldingen på den siste demonstrasjonen, er det forsvarlig å anta at kilden gjorde dette. Videre fulgte vi dette opp med et behov for å validere dataen, slik at vi forsikret at datakvaliteten utgjorde et tilfredsstillende nivå, i henhold til Talend (u.å.). Deretter understreket vi at dataen må transformeres i henhold til bruksformålet. Basert på dette, tilførte informant #1 at standardisering av data kan være kritisk. Avslutningsvis tar vi for oss nødvendigheten ved å berike dataen, dvs. å legge til annen data, eller koble den opp mot relatert informasjon. Under metrikker baserte vi oss på Stedman (2020), derav hans forslag til hvordan metrikker burde utvikles. I denne kilden ble det også diskutert at forretningsmål og det spesifikke brukstilfellet av data burde utgjøre grunnlaget for hva som måles.

Under ferdigheter er det ifølge Kolarov (2020) viktig at ansatte innehar tilstrekkelig kunnskap, slik at man sikrer verdifull ressursbruk. I tillegg forklares det at en generell datakompetanse, hvor selv ansatte som ikke jobber direkte i analyseprosessen, har nok kompetanse om data til å se begrensninger og muligheter ved den. Hvordan dette skal oppnås forklares videre under dimensjonen om ledelse i samme fase. Her blir også tanken understreket av Stedman (2020), som også sier at andre brukere av data ofte bistår i prosessen rundt forberedelse av den til eget bruk. Kolarov sier at det ikke eksisterer noen godt egnede praksiser for å lære opp ansatte, men forklarer videre at ulike tilnærminger som workshops og behovsbasert opplæring kan være nyttige. Vi kombinerte dette med Panetta (2021), som forklarer at man må ha klare svar på hvordan opplæringen kommer de enkelte ansatte til gode, slik at de skal ha noe motivasjon til å lære. Dersom data er en integrert del av virksomheten, vil det nok være utfordrende å lære opp alle som benytter seg av det. I tillegg vil det kreve at de ansatte i faktisk får noe ut av det, sett bort fra kun kunnskap. I ettertid ser vi at dette muligens burde vært utdypet og forklart ytterligere.

Ved beskrivelsen av kultur, baserte vi oss også på Panetta (2021), hvor det forklares at en kultur som verdsetter data, og hvor god datakvalitet ansees som essensielt, kan være viktig i kontekst av dataforberedelse. Vi bygde videre på dette gjennom innsikt fra Amirah Fayek sitert i Stedman (2020). Hun forklarer at bruk av metrikker kan si mye om hvordan datakvalitet verdsettes i virksomheter, og kommer også med forslag til metrikker som vi videre tok for oss i rammeverket. Videre forklarte vi ikke hvordan en slik kultur skal opparbeides, da det i stor grad dekkes i forklaringen av opplæring i datakompetanse under ledelse.

I kapittelet om strategi henviste vi tilbake til strategi under forrige fase, da hovedpoenget er det samme. Dersom ansatte skal vite hvordan data skal forberedes, krever det at de er kjent med virksomhetens strategi, samt mål for prosjektet. Vi følte dermed at dette var dekkende nok. Deretter, under smidighet, bygger vi videre på det som var funnet under samme dimensjon i forrige fase. Dette gjøres basert på både CRISP-DM-modellen og intervju av informant #4. Endringer kan oppstå, noe som vil kreve smidige prosesser for å imøtekomme de. Vi beskrev under datadimensjonen i samme fase, at det eksisterer mange ulike metoder for å oppdage feil, samt reparere data. Endringer kan oppstå som gjør at andre metoder må vurderes, eller så må eventuelt dataen evalueres på nytt.

Oppsummert har vi bidratt til å besvare problemstillingen gjennom å diskutere aktiviteter og faktorer som samlet sett oppfyller kravene til fasen om dataforberedelse. Vi har satt sammen litteratur fra flere kilder samt intervjuer, for å bygge en beskrivelse av aktiviteter som må gjøres for å forberede dataen på en hensiktsmessig måte. I tillegg har

vi forklart faktorer fra de ulike virksomhetsdimensjonene som kan påvirke gjennomførelsen av fasen.

## Modellering

Kravene til modelleringsfasen utgjorde samlet sett at vi måtte ta for oss hvordan virksomheter velger riktig modelleringsteknikk, hvordan de kan benytte dem, hva som må være på plass, samt hva som kan påvirke gjennomførelsen av fasen. Vi konkluderte basert på Wirth & Hipp (2000) og IBM (2021h) at dette innebærer nøye vurderinger av hvilke modelleringsteknikker som best besvarer forretningsbehovene, samt hvilke teknikker som passer den tilgjengelige dataen best. De ulike teknikkene gir ulike resultater, som alle kan passe et gitt behov best. Samtidig er det ifølge IBM (2021h) viktig at dataen samsvarer med modelleringsteknikken, f.eks. med tanke på datakvalitet, datamengde, eller datatype. Vi forklarte ikke de ulike teknikkene, da vi konkluderte med at det var for mange ulike teknikker, samt for teknisk komplekst til at vi ville være i stand til å forklare dette på en hensiktsmessig måte. Essensen er også her, at valgene må forankres i et reelt forretningsbehov. Det var vanskelig å finne relevant litteratur under metrikker i denne fasen. Derfor baserte vi oss på Wirth & Hipp (2000), som foreslo metrikker knyttet til effektivitet i gjennomførelsen av fasen. Dermed forklarte vi at slike metrikker kan hjelpe virksomheter til å bedre forstå prosessen rundt, samtidig som at det kan gi informasjon om de foregående fasene i rammeverket.

Under ferdigheter tok vi for oss like aspekter som i de forrige fasene, altså behovet for ansatte med relevant kunnskap og ferdigheter som kreves for å bruke den aktuelle teknologien. Dette relateres til de ulike modelleringsteknikkene, samt verktøyene som kreves for å gjennomføre dem. Videre forklares dette under dimensjonen om teknologi, hvor vi baserer oss på Wirth & Hipp (2000). De hevder at man krever at virksomheter også besitter den aktuelle teknologien for å gjennomføre modelleringen. Vi skulle ønske vi var i stand til å forklare dette på en mer detaljert måte, og gått mer inn på nøyaktig hva som kreves av teknologi. Til tross for dette, basert på litteraturen vi gjennomgikk, eksisterer det derimot alt for mange modelleringsteknikker og ulike verktøy for ulike behov, til at dette var noe vi kunne gjennomføre på en god måte.

Ved ledelse baserte vi oss på at den viktigste oppgaven her, som ledere også burde utføre, er å overse at modelleringen skjer i samsvar med forretningsbehovet, samt at de ansvarlige for modelleringen har relevant kompetanse innenfor fagfeltet. Under kultur baserte vi oss på Liu et al. (2021), som i rapporten forklarte at virksomheter med kulturer som var kritiske til det utførte arbeidet, samt læringsvillige og kunnskapsdelende, presterte bedre. Resultatet fra denne fasen kan ende opp med å bli det endelige produkt fra prosjektet. Vi fant det derfor hensiktsmessig å inkludere det her. Ansatte som er kritiske, vil i større grad sikre at resultatet faktisk samsvarer med forretningsbehovet. Når det kommer til strategi konkluderte vi med at innholdet var tilsvarende de to foregående fasene, slik at også aktiviteter i denne fasen forankres i reelle forretningsbehov. Da vi beskrev smidighet i denne fasen, baserte vi oss på Wirth & Hipp (2000), som forklarer at dokumentasjon er ekstremt viktig for å sikre oversikt, og dermed bedre tilpasningsevne. De forklarte også at problemer med data ofte kan oppdages i denne fasen av et typisk dataanalyseprosjekt, eller at dataanalytikerne kan få nye ideer til datasett som kan konstrueres. Dette vil kreve at man går tilbake i prosessen, noe som burde legges opp til i slike prosjekter. Dette kan være mer ressurskrevende, men samtidig sørge for at forretningsverden av det endelige resultatet øker.

Oppsummert har vi bidratt til å besvare problemstillingen gjennom å diskutere aktiviteter og faktorer som samlet sett oppfyller kravene til fasen om modellering. Vi baserte oss primært på litteratursøk for å beskrive de ulike aktivitetene og påvirkende faktorene, da vi ikke fikk hentet så mye informasjon om dette ut ifra intervjuene. I denne fasen var det en del ting vi skulle ønske vi hadde kapasitet og kunnskap til å gå mer i detalj på, blant annet ulike modelleringsteknikker og verktøy. Dette kunne gitt en dypere beskrivelse av data-, ferdigheter- og teknologidimensjonen av rammeverket.

## Evaluering

Overordnet gjorde kravene i denne fasen at vi måtte ta for oss aktiviteter som må gjøres, samt faktorer som burde vurderes, under evalueringen av resultatet og prosessen så langt. Under data tok vi for oss hvordan datakvaliteten skulle vurderes, da dette ifølge Chapman (1999) og J.S. Saltz (2021) var essensielt å gjennomføre. Gjennom intervju med informant #3, fikk vi informasjon om at de ofte brukte, eller foreslo til kunder, en kartleggingsprosess de kaller for «*wall of shame*». Her markerer de data som er mangelfull, evt. prosesser eller ansatte som har ført til det. Gjennom dette skal man skille ut avvik, samt bemerke hvor de oppstår. Dette skal være en effektiv metode å sikre datakvalitet på. Samtidig burde det nevnes at navnet muligens er litt misvisende, og at det nok ikke burde brukes for å skammeliggjøre ansatte. Det burde benyttes av ledelsen og evt. andre ansvarlige for å sikre datakvalitet, som kan bidra til å gjøre de nødvendige korrigerende tiltakene. Under metrikker nevnte vi, basert på Ayele (2020), at dette burde benyttes for å evaluere modellresultatene opp mot mål, eller kriterier for prosjektet. Nok en gang vil vi ikke konkretisere noen universelle faktorer som er nyttige for alle prosjekter. Hva som utgjør nyttige metrikker, er avhengig av brukstilfellet og hva de vil oppnå med data.

Basert på Wirth & Hipp (2000), forklarte vi under ferdigheter at virksomheten må besitte ansatte med riktige ferdigheter for å teste modellen opp mot virksomhetens forretningsmål. Vi utformet også en forklaring på hva dette ville utgjøre, basert på denne kilden, i tillegg til Heggernes (2020) og Chapman (1999). Evaluering kan være en fase hvor enormt mange faktorer må tas hensyn til, derfor betraktet vi det som viktig å forklare hva dette innebar her. Under teknologi tok vi for oss at teknologien brukt i prosessen burde evalueres, dvs. om den har vært tilstrekkelig langs prosessen. Kunne f.eks. andre verktøy blitt tatt i bruk for mer gunstige resultater? Var IT-infrastruktur egnet for en smidig gjennomføring av prosjektet? Dette var basert på innsikt fra artikkelen til Ecpu University (u.å.). Vi kombinerte dette med Wirth & Hipp (2000), som forklarer at validering av modellen også er en viktig aktivitet i evalueringsfasen, samt at dette kan kreve ulike teknologier og teknikker som må implementeres basert på behov. Dermed får man en fase hvor både teknologien kontinuerlig i prosessen blir evaluert, samt en forståelse av teknologi som trengs for å utføre selve evalueringen.

Under ledelse tok vi for oss innsikt fra Ayele (2020) og Wirth & Hipp (2000), og kombinerte dette til en forklaring hvor ledelsen må ha ansvar for å definere tydelige kriterier resultatene skal evalueres opp mot, samt sørge for at denne prosessen faktisk blir gjennomført grundig. Det vil basert på dette, også være hensiktsmessig av ledelsen å etablere en plan for mulige fremtidige handlinger basert på resultatet av denne fasen. Under kultur baserte vi oss på innsikt fra informant #1, Liu et al. (2021) samt Anderson (2015). Vi kombinerte innsikten fra disse kildene og beskrev at evalueringsfasen er kritisk, og at dataanalyse ikke trenger å mislykkes mange ganger før ansatte mister troen på bruken av dataanalyser. Som i tidligere steg kan dermed en kritisk, men



læringsvillig og kunnskapsdelende kultur som beskrevet av Liu et al. (2021), være hensiktsmessig. Anderson (2015) sier også at ansatte må gis rom til å eksperimentere og vurdere rundt eget arbeid, slik at de kan lære av eventuelle feil. Vi kombinerte denne innsikten til et forslag om kulturelle elementer virksomheter burde arbeide for å oppnå, for å sikre god gjennomførelse av denne dimensjonen. Ved strategi baserte vi oss på Ayele (2020) og Chapman (1999), og kombinerte innsikten fra deres artikler. De beskriver at man må ta en avgjørelse på hvordan de skal gå videre i denne fasen, dvs. om resultatene er tilfredsstillende nok til å implementeres eller ikke. Hvis en beslutning om å implementere resultatet blir tatt, må man planlegge en strategi for hvordan dette skal gjøres. Vi konkluderte selv med at ledelsen har det overordnede ansvaret for denne avgjørelsen, og følgende at de må sørge for at en plan for implementering blir etablert. Til slutt beskrev vi under smidighet, basert på Chapman (1999) og Wirth & Hipp (2000), at virksomheten må legge opp til at man etter evalueringsfasen kan gå tilbake for å gjøre de nødvendige tiltakene, slik at man ordner opp i svakheter, dersom dette anses som nødvendig. Det kan kreve mer tid og andre ressurser, men dette vil være viktig for å sikre et godt resultat. CRISP-DM legger opp til at dette er en naturlig del av et slikt prosjekt, slik de to kildene vi anvendte her beskriver. Vi anser dette som kritisk også når det kommer til makroanalyser.

Oppsummert har vi i denne fasen svart på kravene til fasen, ved å ta for oss aktiviteter som må gjøres, samt faktorer som bør vurderes under evalueringen av resultatet og den helhetlige prosessen. Igjen, er noen av forklaringene muligens litt generelle, men prosessen rundt evaluering vil være tilnærmet lik når det kommer til makroanalyser og mikroanalyser, så her eksisterer det allerede mer dyptgående kunnskap.

## Implementering

Sett under ett gjorde kravene i denne fasen at vi måtte ta for oss hvilke aktiviteter som kreves for å implementere resultatene, og faktorer som må vurderes i denne sammenhengen. Dette gjelder enten om det er resultat som gir direkte innsikt og kan implementeres i forretningsbeslutninger, eller modeller som f.eks. skal implementeres med andre systemer, samt automatisere analyse av innkommende data.

Innledningsvis i datadimensjonen beskrev vi at dersom de foregående fasene har blitt gjennomført grundig, skal man i teorien ha opparbeidet seg et pålitelig resultat som beslutningstakere kan stole på, samt benytte i beslutningsprosesser. Vi beskrev basert på IBM (2021g) sin forklaring av CRISP-DM, at resultatet kan bestå av primært to ting; en modell som gir direkte innsikt, eller en modell som kan integreres med andre systemer for å automatisere dataanalyseprosesser. Basert på Wirth & Hipp (2000) forstod vi at førstnevnte krever at resultatet visualiseres på en forståelig måte for de som skal benytte seg av det. Sistnevnte kan derimot være mer kompleks, og vi baserte oss på IBM (2021g) for å forklare dette. Her vil det viktigste være at det legges en plan for integrasjonen, for å minimere risikoen av uforutsette hendelser eller problemer som oppstår underveis. Modellene som brukes over lengre tid, må også overvåkes og vedlikeholdes, og nøyaktig hvordan dette skal gjøres bør planlegges.

Under metrikker ville vi berike den forklarte aktiviteten over overvåkning og vedlikehold av implementerte modeller. Vi fant ut at metrikker kan benyttes til å måle hvilke faktorer som indikerer at modellen ikke presterer godt nok, basert på IBM (2021g). Hva som utgjør riktige metrikker, vil dermed avhenge av formålet til modellen, noe vi eksemplifiserer. I tillegg foreslår vi basert på egen kunnskap at metrikker kan benyttes

til å måle effekten som bruken av resultatet har på virksomheten. Eksempelvis kan dette være om det faktisk blir tatt i bruk til å vedta beslutninger, og hvilken effekt dette i så fall har på virksomheten. Under ferdighetsdimensjonen beskriver vi enkelt at dette må utgjøre ferdigheter innen de relevante fagområdene og arbeidet som skal utføres, avhengig av hva dette består av.

Basert på behovet for overvåkning og vedlikehold av modeller, beskriver vi under teknologi at virksomheten må inneha aktuell teknologi for å kunne gjøre dette. Når det kommer til innsikt som skal benyttes, og derfor må visualiseres, beskriver vi basert på Lowe & Metthee (2020) at det finnes mange ulike datavisualiseringsverktøy som dekker ulike behov for dette, som f.eks. Tableau, et verktøy som dekker generelle behov. Noen virksomheter kan derimot ha behov som ikke dekkes av generelle verktøy som dette, og da eksisterer det som Lowe & Metthee (2020) beskriver, mange verktøy som dekker nisjebehov. Under ledelse beskriver vi at implementering, overvåkning og vedlikehold overordnet er ledelsens ansvar, som må sørge for at dette blir gjort. Vi baserer oss i tillegg på teori om endringsledelse fra Thorsvik & Jacobsen (2019) og Todnem (2005), fordi vi så for oss at implementeringen kan føre til endringer eller andre utfordringer som kan møte motstand internt i virksomheten.

Under beskrivelsen av kultur foreslår vi basert på Wirth & Hipp (2000) at det spesifikt må fokuseres på opplæring av de ansatte, da data-drevne makroanalyser er et relativt nytt fagfelt. Vi beriker dette med teori fra Anderson (2015, ss. 214-215). Han beskriver at det er ekstremt viktig at virksomheter, da spesielt ledere, ser verdien i god data og faktisk benytter det i beslutningsprosesser. Ellers sier han at dette kan føre til ringvirkninger basert på at det er lite motiverende å arbeide med noe som ikke benyttes. Dette hevder han at kan føre til problemer i virksomhetens kultur, noe vi følte var viktig å belyse her. Under strategi belyser vi basert på Kang et al. (2022), at selve strategien for implementering bør utformes gjennom en kombinasjon av datavitenskapelig ekspertise og ekspertise rundt data-drevne makroanalyser, for å realisere mest mulig verdi fra resultatet. I tillegg beriker vi dette ved å basere oss på informasjon vi har fått fra intervju, om at dette må forankres i forretningsforståelse, slik at man svarer på reelle forretningsbehov. Til slutt diskuterer vi under smidighet, basert på IBM (2021g), Chapman (1999) og Wirth & Hipp (2000), at integrerte modeller kan slutte å prestere etter etablerte kriterier, noe som kan kreve justeringer i modellene, eller at de eventuelt må avskaffes og bygges opp på nytt. Virksomheten må dermed ha en etablert plan over hva som burde gjøres dersom en modell slutter å prestere etter de etablerte kriteriene. Dette er i samsvar med generelle beskrivelser av CRISP-DM.

Oppsummert har vi her svart på kravene til fasen ved å diskutere aktiviteter som kreves, og faktorer som må vurderes under implementeringen av resultatene. Også her er noen av forklaringene litt generelle, men igjen så vil f.eks. kompetansen som kreves basere seg på verktøy og metoder som benyttes, noe som kan være veldig virksomhetsspesifikt og dermed vanskelig å forklare dekkende i detalj. Sammenlignet med de første tre fasene er det relativt lite nytt å tilføye her, i perspektiv av makroanalyse, utover litteratur som allerede eksisterer om mikroanalyse.

## 7 Konklusjon

Dette kapittelet skal besvare problemstillingen «*Vi vil finne ut hva som kreves teknisk og organisatorisk av et rammeverk for utvikling av data-drevne makroanalyser. Med bakgrunn i dette ønsker vi å forsøke å skissere et rammeverk.*». For å kunne svare på dette velger vi å besvare først prosessen som helhet før vi går videre inn på teknisk og organisatorisk separat.

Gjennom intervju og litteratur utviklet vi et sett med krav over tekniske og organisatoriske faktorer som kreves av et rammeverk for utvikling av data-drevne makroanalyser. På et overordnet nivå utgjorde disse at prosessen må ansees som en helhet. Sammenhengen mellom teknologiske og organisatoriske faktorer er sterk, og de er sterkt avhengig av hverandre. Ved implementering av ny teknologi krever det kompetanse, strategi og ledelse som synliggjør hva hensikten bak endringen er og hva som er ønskelig å oppnå. Man er avhengig av ledere som evner å se både den tekniske og organisatoriske siden av rammeverket og som har troen på at endringen.

I rammeverket vårt fremhever vi at virksomheter må ha strukturerte tilnærminger til de tekniske aktivitetene, samt hvordan disse kan understøttes av riktig teknologi. Spesielt når det kommer innhenting av data, har vi tilført ny kunnskap i form av strukturerte tilnærminger til ekstern data for makroanalyser, som går ut over etablert kunnskap om dataanalyser. Basert på nyere litteratur konkluderer vi med at nye roller, med hovedansvar for ekstern data, er et av hovedelementene som kan gjøre virksomheter kapable til å utføre data-drevne makroanalyser. Bruk av ekstern data kan øke behovet for smidighet, med tanke på overvåkning og vedlikehold av eksterne kilder. Det kan i tillegg øke kompleksiteten av, og mulighetene med data, noe som krever nøye vurderinger og valg basert på forretningsbehov. Vi konkluderer med at mye ulik teknologi kan benyttes, men at valgene må vurderes nøye, og ha en forankring i virksomhetens reelle behov og tilgjengelige ressurser.

Gjennom rammeverket vårt kan vi konkludere med at et grundig forarbeid gjennom opprettelse av forretningsforståelse er kritisk for å konstatere nøyaktig hva virksomheten er ute etter og av hvilken kompleksitet og omfang. Først når dette er tydelig definert kan en begynne prosessen med å utvikle løsningen. En langsiktig plan og strategi kreves for å forberede ansatte på hva som skal skje, hva som kreves av dem og hvorfor dette gjennomføres. Uavhengig om endringen er et resultat av at omgivelsene krever det, eller om virksomheten ønsker å realisere ny verdi i eksisterende data, vil en etablering av forretningsforståelse være elementært. Når det kommer til de siste fasene i rammeverket er det kritisk at riktig og relevant kompetanse eksisterer i virksomheten, slik at man er i stand til å konvertere data til innsikt eller lesbare modeller som får fram nøyaktige resultat. Dette slik at beslutningstakere kan fatte en reflektert beslutning basert på den.

Bacheloroppgavens bidrag til forskning har vært gjennom å undersøke et lite belyst område innen vårt fagområde og se hvordan dette kan være en bidragsyter til utvikling og vekst. Ved å undersøke teori og litteratur rundt data-drevne makroanalyser har vår forskning gjennom et rammeverk synlig- og tydeliggjort hvilke tekniske og organisatoriske faktorer som må ligge til grunn for at en virksomhet skal være i stand til

å utvikle data-drevne makroanalyser. Funnene vi har gjort vil være nyttige for twoday Avento i møte med virksomheter av ulik størrelse og bransje som ønsker å utvikle data-drevne makroanalyser. Når det kommer til videre forskning innenfor temaet, ville det vært svært spennende å se rammeverket bli videre testet og tatt i bruk. På denne måten ville det vært mulig å se hvordan det ville fungert i praksis, eller om det eventuelt trenger justeringer. Ved å sammenligne flere virksomheter gjennom prosessen, kunne man også videreutviklet rammeverket og kanskje avdekket flere dimensjoner og sammenhengen mellom dem.

## 8 Referanser

- Adobe. (2022, august 11). *Types of analytics explained—Descriptive, predictive, prescriptive, and more*. <https://business.adobe.com/blog/basics/descriptive-predictive-prescriptive-analytics-explained>
- Ahmed, N. (2022, august 1). *CTO vs CEO—Who Is Responsible For Business Growth?* <https://enou.co/blog/cto-vs-ceo/>
- Ahmed, P. K. (1998). Culture and climate for innovation. *European Journal of Innovation Management*, 1(1), 30–43. <https://doi.org/10.1108/14601069810199131>
- Al-Mashari, M., & Zairi, M. (1999). BPR implementation process: An analysis of key success and failure factors. *Business Process Management Journal*, 5(1), 87–112. <https://doi.org/10.1108/14637159910249108>
- Anderson, C. (2015). *Creating a Data-Driven Organization: Practical Advice from the Trenches*. O'Reilly Media.
- Aslanova, I. V., & Kulichkina, A. I. (2020). Digital Maturity: Definition and Model. *Advances in Economics, Business and Management Research*, 138, 443–449. <https://doi.org/10.2991/aebmr.k.200502.073>
- AssistKD. (2021, april 12). *Requirements Gathering vs Elicitation*. Assist Knowledge Development. <https://www.assistkd.com/learning-zone/business-alchemists-blog/requirements-gathering-vs-elicitation>
- Aurum, A., & Wohlin, C. (2005). *Engineering and Managing Software Requirements*. Springer.
- Ayele, W. Y. (2020). Adapting CRISP-DM for Idea Mining. *International Journal of Advanced Computer Science and Applications*, 11(6). <https://doi.org/10.14569/IJACSA.2020.0110603>
- Badewi, A. (2022). When frameworks empower their agents: The effect of organizational project management frameworks on the performance of project managers and

- benefits managers in delivering transformation projects successfully. *International Journal of Project Management*, 40(2), 132–141.  
<https://doi.org/10.1016/j.ijproman.2021.10.005>
- Baker, K. (2022, september 14). *The Complete Guide to Business Agility and Why It's Important*. Catalant. <https://catalant.com/business-agility/the-ultimate-guide-to-business-agility-and-why-its-important/>
- Berndtsson, M., Lennerholt, C., Svahn, T., & Larsson, P. (2020). 13 Organizations' Attempts to Become Data-Driven. *International Journal of Business Intelligence Research*, 11(1). <https://doi.org/10.4018/IJBIR.2020010101>
- Beskales, G., Ilyas, I. F., & Golab, L. (2010). Sampling the Repairs of Functional Dependency Violations under Hard Constraint. *Proceedings of the VLDB Endowment*, 3(1), 197–207. <https://doi.org/10.14778/1920841.1920870>
- Bihani, P., & Patil, S. T. (2014). *A Comparative Study of Data Analysis Techniques*. 3(2), 95–101.
- Bohannon, P., Fan, W., Flaster, M., & Rastogi, R. (2005). A cost-based model and effective heuristic for repairing constraints by value modification. *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data*, 143–154. <https://doi.org/10.1145/1066157.1066175>
- Bonthu, S., & Bindu, K. H. (2017). Review of Leading Data Analytics Tools. *International Journal of Engineering & Technology*, 7(3), 10–15.
- Brynjolfsson, E., & McAfee, A. (2017). Artificial Intelligence, For Real. *Harvard Business Review*, BG1704. <https://store.hbr.org/product/artificial-intelligence-for-real/BG1704>
- Buitelaar, R. (2018). *Building the Data-Driven Organization: A Maturity Model and Assessment* [Leiden Institute of Advanced Computer Science].  
<https://theses.liacs.nl/pdf/2017-2018-BuitelaarRuben.pdf>
- Chapman, P. (1999). *The CRISP-DM User Guide*. NCR Systems Engineering Copenhagen.

- Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016). Data Cleaning: Overview and Emerging Challenges. *Proceedings of the 2016 International Conference on Management of Data*, 2201–2206. <https://doi.org/10.1145/2882903.2912574>
- CIPD. (2021, desember 6). *PESTLE Analysis | Factsheets*. CIPD. <https://www.cipd.co.uk/knowledge/strategy/organisational-development/pestle-analysis-factsheet>
- Cote, C. (2021, desember 2). *7 Data Collection Methods in Business Analytics*. Business Insights Blog. <https://online.hbs.edu/blog/post/data-collection-methods>
- Coursera. (2023, februar 23). *What Is Data Analysis? (With Examples)*. Coursera. <https://www.coursera.org/articles/what-is-data-analysis-with-examples>
- Data Science Association. (2020, august 10). *About Data Science*. <https://web.archive.org/web/20200810114002/http://www.datascienceassn.org/about-data-science>
- Dhar, V. (2013, desember). *Data Science and Prediction*. <https://cacm.acm.org/magazines/2013/12/169933-data-science-and-prediction/abstract>
- Digitaliseringsdirektoratet. (2022, august 30). *Felles datakatalog*. <https://data.norge.no>
- Epci University. (u.å.). *What is IT Infrastructure?*
- gov.uk. (2021, juni 24). *Meet the data quality dimensions*. Government Data Quality Hub. <https://www.gov.uk/government/news/meet-the-data-quality-dimensions>
- Gray, D. (2017, juli 22). What is Culture Mapping and why should you care? *The XPLANE Collection*. <https://medium.com/the-xplane-collection/what-is-culture-mapping-and-why-should-you-care-912f99c4de98>
- Griffin, J. (2008). The Role of the Chief Data Officer. *DM Review New York*, 18(2). <https://www.proquest.com/docview/214672505?pq-origsite=gscholar&fromopenview=true>
- Halper, F., & Stodder, D. (2017). *What it Takes to Be Data-Driven*. TDWI. [https://media.bitpipe.com/io\\_14x/io\\_141315/item\\_1674359/TDWI\\_BPReport\\_Q417.pdf](https://media.bitpipe.com/io_14x/io_141315/item_1674359/TDWI_BPReport_Q417.pdf)

- Heggernes, T. A. (2020). Digital forretningsforståelse. I *Digital forretningsforståelse, Fra store data til små biter* (3.). Fagbokforlaget.
- Heinrich, B., Hristova, D., Klier, M., Schiller, A., & Szubartowicz, M. (2018). Requirements for Data Quality Metrics. *Journal of Data and Information Quality*, 9(2), 1–32. <https://doi.org/10.1145/3148238>
- Helo, P., & Hao, Y. (2021). Artificial intelligence in operations management and supply chain management: An exploratory case study. *Product Planning & Control*, 33(1), 1–18. <https://www.doi.org/10.1080/09537287.2021.1882690>
- Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (2016). *The age of analytics: Competing in a data-driven world* (s. 126). McKinsey Global Institute.
- IBM. (u.å.). *What is a data architecture?* Hentet 6. april 2023, fra <https://www.ibm.com/topics/data-architecture>
- IBM. (2011). *IBM SPSS Modeler CRISP-DM Guide*. IBM. [https://inseaddataanalytics.github.io/INSEADAnalytics/CRISP\\_DM.pdf](https://inseaddataanalytics.github.io/INSEADAnalytics/CRISP_DM.pdf)
- IBM. (2021a, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview>
- IBM. (2021b, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=understanding-business-overview>
- IBM. (2021c, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=understanding-data-overview>
- IBM. (2021d, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=preparation-data-overview>
- IBM. (2021e, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=modeling-overview>
- IBM. (2021f, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=evaluation-overview>
- IBM. (2021g, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=evaluation-overview>



- IBM. (2021h, august 17). *IBM Documentation*. <https://www.ibm.com/docs/en/spss-modeler/saas?topic=techniques-choosing-right-modeling>
- Innovasjon Norge. (2021). *Innovasjonstalen 2021 Statusrapport: Koronakrisens påvirkning på næringslivet*. [https://www.innovasjon norge.no/globalassets/0-innovasjon norge.no/om-innovasjon-norge/nyheter/innovasjonstalen-rapport\\_enkeltsider.pdf](https://www.innovasjon norge.no/globalassets/0-innovasjon norge.no/om-innovasjon-norge/nyheter/innovasjonstalen-rapport_enkeltsider.pdf)
- Jin, Z. (2018). *Environment Modeling-Based Requirements Engineering for Software Intensive Systems*. Elsevier.  
<https://www.sciencedirect.com/science/article/pii/B9780128019542000029>
- Kang, H.-G., Jang, G.-Y., & Choi, M. (2022). Data-driven Value-enhancing Strategies: How to Increase Firm Value Using Data Science. *Asia Pacific Journal of Information Systems*, 32(3), 477–495.  
<https://doi.org/10.14329/apjis.2022.32.3.477>
- Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research – Moving away from the “What” towards the “Why”. *International Journal of Information Management*, 54, 102205.  
<https://doi.org/10.1016/j.ijinfomgt.2020.102205>
- Kherdekar, V. A., & Metkewar, P. S. (2016). A Technical Comprehensive Survey of ETL Tools. *International Journal of Applied Engineering Research*, 11(04), 2557–2559.  
<https://doi.org/10.37622/IJAER/11.4.2016.2557-2559>
- Korolov, M. (2020, august). *Why more employees need data literacy skills | TechTarget*. Business Analytics.  
<https://www.techtarget.com/searchbusinessanalytics/feature/Why-more-employees-need-data-literacy-skills>
- Legner, C., Krasikov, P., & Eurich, M. (2022). *Unleashing the Potential of External Data: A DSR-based Approach to Data Sourcing*. 16.  
[https://www.researchgate.net/publication/360603471\\_Unleashing\\_the\\_Potential\\_of\\_External\\_Data\\_A\\_DSR-based\\_Approach\\_to\\_Data\\_Sourcing](https://www.researchgate.net/publication/360603471_Unleashing_the_Potential_of_External_Data_A_DSR-based_Approach_to_Data_Sourcing)

- Lemieux, A. (u.å.). *Academic Guides: Theories and Frameworks: Introduction*. Hentet 12. april 2023, fra <https://academicguides.waldenu.edu/library/theory/introduction>
- Lim, W. M., Chin, M. W. C., Ee, Y. S., Fung, C. Y., Giang, C. S., Heng, K. S., Kong, M. L. F., Lim, A. S. S., Lim, B. C. Y. L., Lim, R. T. H., Lim, T. Y., Ling, C. C., Mandrinos, S., Nwobodo, S., Phang, C. S. C., She, L., Sim, C. H., Su, S. I., Wee, G. W. E., & Weissmann, M. A. (2022). What is at stake in a War? A prospective evaluation of the Ukraine and Russia conflict for business and society. *Global Business and Organizational Excellence*, 41(6), 1–46. <https://doi.org/10.1002/joe.22162>
- Liu, G., Tsui, E., & Kianto, A. (2021). Knowledge-friendly organisational culture and performance: A meta-analysis | Elsevier Enhanced Reader. *Journal of Business Research*, 134(2021), 738–753. <https://doi.org/10.1016/j.jbusres.2021.05.048>
- Lowe, J., & Matthee, M. (2020). Requirements of Data Visualisation Tools to Analyse Big Data: A Structured Literature Review. *Responsible Design, Implementation and Use of Information and Communication Technology*, 469–480. [https://link.springer.com/chapter/10.1007/978-3-030-44999-5\\_39](https://link.springer.com/chapter/10.1007/978-3-030-44999-5_39)
- Lycett, M. (2013). «Datafication»: Making sense of (big) data in a complex world. *European Journal of Information Systems*, 22, 381–386. <https://doi.org/10.1057/ejis.2013.10>
- Mabert, V. A., Soni, A., & Venkataramanan, M. A. (2003). Enterprise resource planning: Managing the implementation process. *European Journal of Operational Research*, 146(2), 302–314. [https://doi.org/10.1016/S0377-2217\(02\)00551-9](https://doi.org/10.1016/S0377-2217(02)00551-9)
- Magnussen, N. (2020, mai 28). En enkel guide til digital markedsføringsstrategi. *DIG2100 (2103)*. <https://dig2100.no/en-enkel-guide-til-digital-markedsforingsstrategi/>
- Martin, S. (2020, juli 17). *How Data Science will Impact Future of Businesses?* Medium. <https://towardsdatascience.com/how-data-science-will-impact-future-of-businesses-7f11f5699c4d>
- Martinez-Caro, E., Cegarra-Navarro, J. G., & Alfonso-Ruiz, F. J. (2020). Digital technologies and firm performance: The role og digital organisational culture.

- Technological Forecasting & Social Change*, 154.  
<https://doi.org/10.1016/j.techfore.2020.119962>
- Menshikov, S. (2017, mars 26). *Top 6 Ways To Assess Employee Skills And Competencies*. ELearning Industry. <https://elearningindustry.com/top-6-ways-assess-employee-skills-and-competencies>
- Miloslavskaya, N., & Tolstoy, A. (2016). *Big Data, Fast Data and Data Lake Concepts*. 88, 300–305. <https://doi.org/10.1016/j.procs.2016.07.439>
- Mintzberg, H. (1979). *The Structuring of Organizations*. Pearson.
- Mysore, D., Khupat, S., & Jain, S. (2013, oktober 14). *Archived | Understanding the architectural layers of a big data solution—IBM Developer*. <https://developer.ibm.com/articles/bd-archpatterns3/>
- Namugenyi, C., Nimmagadda, S. L., & Reiners, T. (2019). Design of a SWOT Analysis Model and its Evaluation in Diverse Digital Business Ecosystem Contexts. *Procedia Computer Science*, 159, 1145–1154. <https://doi.org/10.1016/j.procs.2019.09.283>
- Narasimman, P. (2023, januar 24). *Data Science Team Structure [How to Build + Best Practices]*. <https://www.knowledgehut.com/blog/data-science/data-science-team-structure>
- Nikiforova, A. (2020). Definition and Evaluation of Data Quality: User-Oriented Data Object-Driven Approach to Data Quality Assessment. *Baltic Journal of Modern Computing*, 8(3). <https://doi.org/10.22364/bjmc.2020.8.3.02>
- Nucleus Research. (2014). *Analytics Pays Back \$13.01 for Every Dollar Spent*. Nucleus Research. <https://newintelligence.ca/wp-content/uploads/2015/02/Nucleus-research-ROI-Sep-2014.pdf>
- O’Neil, C., & Schutt, R. (u.å.). *1. Introduction: What Is Data Science? - Doing Data Science [Book]*. Hentet 21. mars 2023, fra <https://www.oreilly.com/library/view/doing-data-science/9781449363871/ch01.html>
- Ot, A. (2023, februar 10). *What Is Raw Data? Definition, Examples & Processing Steps*. Datamation. <https://www.datamation.com/big-data/raw-data/>

- Panetta, K. (2021, august 26). *A Data Literacy Guide For D&A Leaders*. Gartner.  
<https://www.gartner.com/smarterwithgartner/a-data-and-analytics-leaders-guide-to-data-literacy>
- Panoply. (u.å.). *Cloud Data Warehouse vs Traditional Data Warehouse Concepts*.  
Panoply. Hentet 9. april 2023, fra <https://panoply.io/data-warehouse-guide/data-warehouse-concepts-traditional-vs-cloud/>
- Patil, D. J., & Mason, H. (2015). *Data Driven*. O'Reilly Media, Inc.  
[https://books.google.no/books?hl=no&lr=&id=yOFwBgAAQBAJ&oi=fnd&pg=PP1&dq=Patil,+D.+J.,+%26+Mason,+H.+\(2015\).+Data+Driven.+O%E2%80%99Reilly+Media,+Inc.&ots=x27Jj2wZxl&sig=bgy1RtdJphEZm26d3JH5B5THil8&redir\\_esc=y#v=onepage&q&f=false](https://books.google.no/books?hl=no&lr=&id=yOFwBgAAQBAJ&oi=fnd&pg=PP1&dq=Patil,+D.+J.,+%26+Mason,+H.+(2015).+Data+Driven.+O%E2%80%99Reilly+Media,+Inc.&ots=x27Jj2wZxl&sig=bgy1RtdJphEZm26d3JH5B5THil8&redir_esc=y#v=onepage&q&f=false)
- Peppard, J., & Ward, J. (2007). *Managing the Realization of Business Benefits from IT Investments*. University of Minnesota.
- Piatetsky, G. (2014, oktober 28). CRISP-DM, still the top methodology for analytics, data mining, or data science projects. *KDnuggets*. <https://www.kdnuggets.com/crisp-dm-still-the-top-methodology-for-analytics-data-mining-or-data-science-projects.html>
- Porter, M. E. (2008). The Five Competitive Forces That Shape Strategy. *Harvard Business Review*, 86(1), 78–93.
- Pratt, M. K. (2021, august 29). *How to Structure and Manage a Data Science Team*. Business Analytics.  
<https://www.techtarget.com/searchbusinessanalytics/feature/How-to-structure-and-manage-a-data-science-team>
- Radovilsky, Z., Hegde, V., Acharya, A., & Uma, U. (2018). Skills Requirements of Business Data Analytics and Data Science Jobs: A Comparative Analysis. *Journal of Supply Chain and Operations Management*, 16(1).  
<https://www.csupom.com/uploads/1/1/4/8/114895679/v16n1p5.pdf>
- Ransbotham, S., Kiron, D., & Prentice, P. K. (2016). Beyond the Hype: The Hard Work Behind Analytics Success. *MIT Sloan Management Review*.

- <https://www.semanticscholar.org/paper/Beyond-the-hype%3A-The-hard-work-behind-analytics-Ransbotham-Kiron/874a887feee27a13268230c7ce0ad87a256c9264>
- Reding, M. (2021, oktober 6). *What is a PESTLE Analysis?* CPD Online College. <https://cpdonline.co.uk/knowledge-base/business/pestle-analysis/>
- Regjeringen. (2020). *Nasjonal strategi for kunstig intelligens*. Kommunal- og moderniseringsdepartementet. <https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf>
- Rossen, E., & Nätt, T. H. (2022). API. I *Store norske leksikon*. <https://snl.no/API>
- Sabbaghi, A., & Vaidyanathan, G. (2004). *SWOT Analysis and Theory of Constraint in Information Technology Projects*.
- Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. <https://doi.org/10.1109/CTS.2013.6567202>
- Saltz, J. (2022, februar 5). CRISP-DM is Still the Most Popular Framework for Executing Data Science Projects. *Data Science Process Alliance*. <https://www.datascience-pm.com/crisp-dm-still-most-popular/>
- Saltz, J. S. (2021). CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps. *2021 IEEE International Conference on Big Data (Big Data)*, 2337–2344. <https://doi.org/10.1109/BigData52589.2021.9671634>
- SAP. (u.å.). *Hva er et datavarehus? | Definisjon, komponenter, arkitektur | SAP Insights*. SAP. Hentet 9. april 2023, fra <https://www.sap.com/norway/insights/what-is-a-data-warehouse.html>
- Saras. (2022, juli 26). *How To Choose The Right ETL Tool For Your Business*. <https://sarasanalytics.com/blog/how-to-choose-the-right-etl-tool/>
- Schmeer, K. (2000). *Stakeholder Analysis Guidelines*. [https://www.researchgate.net/publication/265021546\\_Stakeholder\\_Analysis\\_Guidelines](https://www.researchgate.net/publication/265021546_Stakeholder_Analysis_Guidelines)

- Sherman, R. (2015). *Business Intelligence Guidebook*. Elsevier.  
<https://www.sciencedirect.com/book/9780124114616/business-intelligence-guidebook>
- Sidi, F., Shariat Panahy, P. H., Affendey, L. S., Jabar, M. A., Ibrahim, H., & Mustapha, A. (2012). *Data quality: A survey of data quality dimensions*. 300–304.  
<https://doi.org/10.1109/InfRKM.2012.6204995>
- SINTEF. (u.å.). *Hva er kunstig intelligens?* SINTEF. Hentet 13. mars 2023, fra  
<https://www.sintef.no/fagomrader/kunstig-intelligens/hva-er-kunstig-intelligens/>
- Smith, L. W. (2000, juli 9). *Stakeholder analysis—A pivotal practice of successful projects*. Stakeholder Engagement, Houston, Texas.  
<https://www.pmi.org/learning/library/stakeholder-analysis-pivotal-practice-projects-8905>
- Sporse, T. (2021, november 24). *På vei mot omforente metrikker i IT-bransjen*.  
<https://blogg.knowit.no/solutions-no/på-vei-mot-omforente-metrikker-i-it-bransjen>
- Sreemathy, J., Brindha, R., Selva, M. N., Suvekha, N., Karthick, R. N., & Praveennandha, M. (2021). *Overview of ETL Tools and Talend-Data Integration*. 1650–1654.  
<https://www.doi.org/10.1109/ICACCS51430.2021.9441984>
- Stedman, C. (2020, februar). *Data governance metrics: Data quality, data literacy and more* | TechTarget. Data Management.  
<https://www.techtarget.com/searchdatamanagement/feature/Data-governance-metrics-Data-quality-data-literacy-and-more>
- Stedman, C. (2022, januar). *What is Data Cleansing (Data Cleaning, Data Scrubbing)?* TechTarget. <https://www.techtarget.com/searchdatamanagement/definition/data-scrubbing>
- Store Norske Leksikon. (2023). *Bransje*. <https://snl.no/bransje>
- Talend. (u.å.). *What is Data Preparation? Processes and Example*. Talend - A Leader in Data Integration & Data Integrity. Hentet 21. april 2023, fra  
<https://www.talend.com/resources/what-is-data-preparation/>

- Terry Bahill, A., & Henderson, S. J. (2005). Requirements development, verification, and validation exhibited in famous failures. *Systems Engineering*, 8(1), 1–14.  
<https://doi.org/10.1002/sys.20017>
- Thorsvik, J., & Jacobsen, D. I. (2019). Hvordan organisasjoner fungerer. I *Hvordan organisasjoner fungerer* (5., s. 372–378, 391–392). Fagbokforlaget.
- Todnem, R. (2005, desember 1). *Organisational change management: A critical review*.  
<https://doi.org/10.1080/14697010500359250>
- TOGAF. (u.å.). *The TOGAF Standard, Version 9.2—Definitions*.
- twoday. (u.å.). *Om oss—Twoday*. Hentet 6. mars 2023, fra <https://www.twoday.no/om-twoday>
- Twoday Avento*. (u.å.). Hentet 6. mars 2023, fra <https://www.twoday.no/twoday-avento>
- Vikøren, B. M., & Pihl, R. (2023). SWOT-analyse. I *Store norske leksikon*.  
<https://snl.no/SWOT-analyse>
- Visure. (u.å.). Kravanalyse. *Visure Solutions*. Hentet 18. mars 2023, fra <http://visuresolutions.com/no/sporbarhetsveiledning-for-kravstyring/kravanalyse/>
- Volk, M., Jamous, N., & Turowski, K. (2017, august 10). *Ask the Right Questions: Requirements Engineering for the Execution of Big Data Projects*. Twenty-Third Americas Conference on Information Systems, Boston.  
[https://www.researchgate.net/profile/Matthias-Volk-4/publication/338740292\\_Ask\\_the\\_Right\\_Questions\\_Requirements\\_Engineering\\_for\\_the\\_Execution\\_of\\_Big\\_Data\\_Projects\\_Full\\_Paper/links/5e281a7f4585150ee77868b3/Ask-the-Right-Questions-Requirements-Engineering-for-the-Execution-of-Big-Data-Projects-Full-Paper.pdf](https://www.researchgate.net/profile/Matthias-Volk-4/publication/338740292_Ask_the_Right_Questions_Requirements_Engineering_for_the_Execution_of_Big_Data_Projects_Full_Paper/links/5e281a7f4585150ee77868b3/Ask-the-Right-Questions-Requirements-Engineering-for-the-Execution-of-Big-Data-Projects-Full-Paper.pdf)
- Wahono, R. S. (2003). Analyzing Requirements Engineering Problems. *Proceedings of the IECI Japan Workshop 2003*, 55–58. <https://ilmukomputer.org/wp-content/uploads/2007/10/romi-ijw2003.pdf>
- Wang, J., Kraska, T., Franklin, M. J., & Feng, J. (2012). CrowdER: Crowdsourcing entity resolution. *Proceedings of the VLDB Endowment*, 5(11), 1483–1494.  
<https://doi.org/10.14778/2350229.2350263>

- Wieczorkowski, J., & Polak, P. (2014). Big data: Three-aspect approach. *Online Journal of Applied Knowledge Management, 2*, 2014.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 1*, 29–39.
- Yakout, M., Elmagarmid, A. K., Neville, J., Ouzzani, M., & Ilyas, I. F. (2011). Guided data repair. *Proceedings of the VLDB Endowment, 4(5)*, 279–289.  
<https://doi.org/10.14778/1952376.1952378>
- Yüksel, I. (2012). Developing a Multi-Criteria Decision Making Model for PESTEL Analysis. *International Journal of Business and Management, 7(24)*, 52–66.  
<https://www.doi.org/10.5539/ijbm.v7n24p52>
- Zhu, J. Y., Tang, B., & Li, V. O. K. (2019). A five-layer architecture for big data processing and analytics. *International Journal of Big Data Intelligence, 6(1)*, 38–49.
- Aaser, M., & McElhaney, D. (2021, mars 2). *Harnessing the power of external data*.  
<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/harnessing-the-power-of-external-data>



# 9 Vedlegg

## 9.1 Vedlegg 1: Resultat interessentanalyse

### **Interne interessenter**

**Toppledelse** - Toppledelsen er det øverste nivået av ledelsen i en virksomhet, og har ansvaret for den daglige driften av virksomheten. Dette innbefatter tradisjonelt sett et stort fokus på økonomi og resultater. Toppledelsen i virksomheter kan bruke rammeverket til å ta strategiske beslutninger om virksomhetens fremtidige retning, basert på innsikt rammeverket gir. Et av formålene med utviklingen av rammeverket er at sluttproduktet vil fungere som en veileder for ledere av virksomheter som ønsker å ta i bruk makroanalyser. Toppledelsen vil ha høy interesse i et rammeverk som dette, da økt bruk av makroanalyser for å danne bedre beslutningsgrunnlag kan gi store gevinster. Det vil også være viktig å avdekke deres krav, da det i stor grad er ment for å bli tatt i bruk av toppledelsen. Vi betrakter dermed toppledelsens makt som stor.

**Dataanalytikere** - Dataanalytikere jobber med å behandle og analysere data virksomheten besitter, med mål om å utarbeide verdifull innsikt for virksomheten. Det er disse som vil jobbe direkte med å utføre makroanalyser, og vil dermed stille en rekke krav til ulike ressurser de krever for å utføre arbeidet sitt. Dataanalytikere vil ha høy interesse av rammeverket vårt, da det blant annet kan føre til endringer som understøtter deres arbeid. De vil også ha høy makt, da de innehar mye kunnskap om de tekniske faktorene som må inngå i rammeverket.

**IT-ansvarlige** - Under IT-ansvarlig betrakter vi de som har ansvaret med å styre og drifte virksomhetens helhetlige IT-arkitektur, det vil si både maskin- og programvare. De IT-ansvarlige vil med stor sannsynlighet bli påvirket av tiltak gjort etter rammeverket vårt. IT-ansvarlige vil også kunne ha bruk for et slik rammeverk som en veileder for hva som kreves av arbeidet deres. IT-ansvarlige vil dermed ha noe høy interesse i prosjektet. De vil også ha noe høy makt.

**Avdelingsledere** - Avdelingsledere har det overordnede administrative ansvaret overfor deres avdeling. Typisk ansvar er ofte økonomiske og faglige, samt også over avdelingens ansatte. Avdelingsledere vil kunne bruke rammeverket til å bedre forstå hvordan deres avdeling blir påvirket, hva som kreves av dem, samt til å fatte beslutninger innenfor deres respektive ansvarsområder. Avdelingsledere vil ha noe høy interesse i prosjektet, samt noe høy makt.

**Øvrige ansatte** - Under øvrige ansatte betrakter vi de ansatte i virksomheter som ikke har vært dekket i interessentanalysen så langt. Disse vil kunne bli påvirket av endringer rammeverket vårt kan frembringe i virksomheter. Vi vil betrakte deres interesse som middels, og makt som noe lav.

## Eksterne interessenter

**Eksterne konsulenter** - Disse interessentene kan bruke rammeverket til å bedre tilby konsulenttjenester til virksomheter som kan ha behov for å utvikle bruken av data-drevne makroanalyser. De vil derfor ha høy interesse i utviklingsprosjektet. De vil også ha høy makt, da rammeverket i stor grad er ment å kunne benyttes av eksterne konsulenter.

**Kunder eller klienter** - Disse interessentene kan indirekte dra nytte av virksomhetens bruk av rammeverket, dersom det resulterer i høyere kunde verdi. Kunder og klienter er i liten grad direkte involvert i virksomhetens interne prosesser, og vi anser dermed interessen for rammeverket som lav. Kunder og klienter vil i mindre grad stille krav til interne tekniske og organisatoriske faktorer hos virksomheter, så vi anser også makten deres som lav.

**Myndigheter** - Myndigheter stiller krav til virksomheter som opererer under deres myndighetsområder. Sett i kontekst av makroanalyser, kan dette være lover om behandling av data, lover om lagring av sensitiv data, lover om bruk av teknologi, etc. Myndigheter vil ha lav interesse i prosjektet, men noe høy grad av makt.

**Investorer** - Investorer er personer eller firma som investerer i virksomheten, med mål om å oppnå profitt. Investorer har stor interesse i at virksomheten de investerer i oppnår gode resultater, men mindre interesse i alt som foregår internt. Vi anser derfor interessen deres i et rammeverk som virksomheter kan ta i bruk som lav. Investorer har typisk lite direkte makt over virksomhetens beslutninger, og vi anser dermed også deres makt som lav.

Interessentanalysen vår resulterte i følgende oversikt:

**Tabell 5 Resultat av interessentanalyse**

| Interessent           | Makt | Interesse | Totalt |
|-----------------------|------|-----------|--------|
| Toppledelse           | 5    | 5         | 10     |
| Dataanalytikere       | 5    | 5         | 10     |
| IT-ansvarlige         | 4    | 4         | 8      |
| Avdelingsledere       | 4    | 4         | 8      |
| Øvrige ansatte        | 3    | 2         | 5      |
| Eksterne konsulenter  | 5    | 5         | 10     |
| Kunder eller klienter | 1    | 1         | 2      |
| Myndigheter           | 1    | 4         | 5      |
| Investorer            | 1    | 1         | 2      |

1 = lav. 2 = noe lav. 3 = middels. 4 = noe høy. 5 = høy.

Av denne kan vi se at toppledelse, dataanalytikere, IT-ansvarlige, avdelingsledere, og eksterne konsulenter scoret høyest totalt, og vi valgte derfor å betrakte dem videre i prosessen. Vi valgte å ikke inkludere myndigheter, øvrige ansatte, og kunder eller klienter videre i prosessen. Til tross for at vi anser myndighetenes makt som noe høy, valgte vi å ikke undersøke deres krav videre. Vi ønsket derimot gjennom rammeverket å belyse at virksomheter må sørge for at lover og reglementer blir overholdt. Nøyaktig hvilke lover virksomheter må forholde seg til, vil i stor grad variere. Vi anså det derfor som lite hensiktsmessig å analysere deres krav i dybden videre i utviklingen av dette

rammeverket. Vi valgte også bort øvrige ansatte og kunder eller klienter, basert både på analysen og at informasjonen vi kan hente fra disse, i de fleste tilfeller vil være mindre nyttig i utviklingen av rammeverket, enn informasjonen vi kan hente fra de utvalgte interessentene.

### Formål med å undersøke utvalgte interessenters krav

For å bedre være i stand til å utvikle hensiktsmessige spørsmål for kravfremkalling, anså vi det som nyttig å etablere formålet vårt med å undersøke de utvalgte interessentene. Gjennom dette kan vi avgrense spørsmålene vi utvikler til de ulike informantene vi har tilgjengelig hos twoday Avento.

**Toppledelsen** - Ved utviklingen av rammeverket vil det være ekstremt viktig å undersøke og betrakte toppledelsens krav til sluttproduktet. Vi har tidligere kommet fram til at ledelsen må ta både tekniske og organisatoriske faktorer i betraktning ved utviklingen av strategier for gjennomførelsen av digitaliseringsprosjekter, for å øke sannsynligheten for å lykkes. Toppledelsen kan dermed ha stor nytte av et slikt rammeverk, dersom de ønsker å utvikle bruken av data-drevne makroanalyser innad virksomhet deres. For å øke verdien av rammeverket ovenfor denne interessenten, vil det være viktig å avdekke deres krav til rammeverket.

**Dataanalytikere** - For at dataanalytikerne skal kunne utføre arbeidet deres, kreves det at ulike tekniske og organisatoriske faktorer er på plass. Det vil dermed være essensielt å avdekke dataanalytikernes krav til rammeverket, da dette er ment til å understøtte deres potensielle arbeidsoppgaver. Gjennom å undersøke deres krav til rammeverket, vil vi kunne avdekke både tekniske og organisatoriske krav de stiller. Tekniske krav kan for eksempel være krav til data og verktøy, mens organisatoriske krav potensielt kan være støtte fra ledelsen og krav til kompetanse i teamene som settes sammen.

**IT-ansvarlige** - IT-ansvarlige kan ha mye verdifull innsikt i hva som kreves av rammeverket vi utvikler, i henhold til hvordan vi skal forholde oss til virksomheters IT-arkitektur. Det vil derfor være viktig å undersøke disse i utviklingsprosessen vår.

**Avdelingslederne** - For at endringer skal lykkes i virksomheter, må de ikke kun være innforstått av toppledelsen, men også avdelingsledelse og øvrige involverte ansatte. Det kan derfor være viktig å undersøke avdelingslederens krav da det kan bidra til økt forståelse av deres behov av produktet vårt. Dette kan øke sannsynligheten for produktets nytteverdi for virksomheters som helhet, dersom det oppfyller avdelingslederens krav til å ta det i bruk.

**Eksterne konsulenter** - Rammeverket er i stor grad ment til å kunne benyttes av eksterne konsulenter som arbeider med å utvikle bruken av data-drevne makroanalyser i virksomheter. For å sikre at denne interessentgruppen kan hente verdi i å ta i bruk rammeverket, vil det være viktig å forstå deres krav til produktet.

## 9.2 Vedlegg 2: Samtykkeerklæring

# Vil du delta i forskningsprosjektet

## «*Digital forretningsutvikling*»?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å kunne studere anvendelse av IT og hvordan dette kan skape gevinster for virksomheten. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

### **Formål**

Denne oppgaven er en bacheloroppgave i studiet Bachelor i Digital forretningsutvikling ved Institutt for datateknologi og informatikk ved NTNU. Vi vil forsøke å belyse et tema tilhørende den overordnede problemstillingen om hvordan anvendelse av IT på ulike måte kan skape gevinster for virksomheten.

Overordnet problemstilling er:

*«Vi vil finne ut hva som kreves teknisk og organisatorisk av et rammeverk for utvikling av data-drevne makroanalyser.»*

### **Hvem er ansvarlig for forskningsprosjektet?**

NTNU er ansvarlig for prosjektet.

### **Hvorfor får du spørsmål om å delta?**

I forbindelse med vår bacheloroppgave, har vi valgt å gå for en kvalitativ forskningsmetode. På bakgrunn av dette så ønsker vi å komme i kontakt med tre til fire informanter i twoday Avento, som kan inneha relevant kunnskap for oss.

### **Hva innebærer det for deg å delta?**

Hovedsakelig ønsker vi å hente inn informasjon gjennom intervju, men på grunn av plasseringsforskjeller, kan det være aktuelt gjennomføre intervju digital eller via e-post. Dersom du ønsker å delta i prosjektet innebærer det at du svarer på mellom 5-8 spørsmål. Det vil ta deg max 30 minutter. Spørsmålene handler om innhenting, forvaltning og kvalitet av data, implementering og eventuelle endring i bedrift. Dine svar fra spørreskjemaet blir registrert elektronisk og i oppgaven vil du være anonym og omtales som *«informant #»*

### **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.

- *Ved behandlingsansvarlig institusjon vil prosjektgruppe og veileder ha tilgang.*
- *Lydopptak ved intervjuer vil lagres på sikret nettverk/digital plattform der NTNU har databehandleravtale*
- *Spørreundersøkelse gjennomføres og lagres på digital plattform der NTNU har databehandleravtale*

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er 22.05.2023. *Personopplysninger og lydopptak slettes ved prosjektslutt.*

### **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 NTNU 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:

- NTNU ved Leif Erik Opland (leif.e.opland@ntnu.no).
- Vårt personvernombud: Thomas Helgesen.
- NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@sikt.no) eller telefon: 53 21 15 00.

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@sikt.no) eller på telefon: 53 21 15 00.

Med vennlig hilsen

Prosjektansvarlige  
*Beate Drønnen, John-Kevin Sydow Smådal og Thomas Lande*

Veileder  
*Thomas Østerlie*

---

## Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Digital forretningsutvikling*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju
- å delta i digitalt intervju eller intervju pr e-post

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet 1.6.2023

---

(Signert av prosjektdeltaker, dato)

## 9.3 Vedlegg 3: Intervjuguide #1

Før oppstart

- Presenterer gruppen og prosjektet og formålet ved intervjuet
- Går gjennom informasjonsskriv og innhenter skriftlig samtykke til deltagelse

| <b>Tema</b>                                     | <b>Spørsmål</b>  |
|---|--|
| <b>Intro</b>                                    | <ol style="list-style-type: none"><li>1. Navn og alder</li><li>2. Hva er din rolle/stilling?</li><li>3. Kan du fortelle om din erfaring/ tidligere utdanning/bakgrunn?</li><li>4. Hvor lenge har du vært ansatt/leder i bedriften?</li></ol>   |
| <b>Spørsmål om data og teknologi</b>            | <ol style="list-style-type: none"><li>1. Hva kreves av data for at bedrifter kan ta i bruk data-drevne prediksjoner?</li><li>2. Hva er typiske problemer som oppstår mtp. datakvalitet?</li><li>3. Hvordan hjelper dere bedrifter sikre god datakvalitet?</li><li>4. Hvordan henter virksomheter typisk inn data i dag? Er det vanlig for virksomheter å benytte seg av tredjepartsdata?</li><li>5. Hvordan kan man best mulig prioritere eksterne datakilder? (f.eks. basert på relevante makrofaktorer, evt. bransje og næring)</li></ol>  |
| <b>Spørsmål om virksomhet</b>                   | <ol style="list-style-type: none"><li>1. Hva burde vektlegges av bedrifter når det kommer til digitalisering?</li><li>2. Hva er den vanligste grunnen til at bedrifter ønsker å digitaliseres?</li><li>3. Opplever dere at det er et stort fokus fra ledelsen i virksomheter på å bli datadrevet?</li><li>4. Opplever dere økt motstand mot digitalisering, gitt utviklingen de siste årene? Teknostress, etc.</li></ol>   |
| <b>Spørsmål om fremgangsmåte i virksomheten</b> | <ol style="list-style-type: none"><li>1. Hvordan går dere typisk frem når dere inngår samarbeid med virksomheter? Hva kartlegges i innledende fase?</li><li>2. Ser dere en tydelig likhet mellom bedriftene som ønsker digitalisering?</li><li>3. Hvor langt opplever dere at bedrifter er kommet i en digitaliseringsprosess, når de henvender seg til dere? Evt. Hvorfor?</li><li>4. Ser du noen tydelige trender i hva bedriftene etterspør med tanke på teknologi?</li><li>5. Hvor langt opplever dere at bedrifter ønsker å strekke seg ressursmessig, i forhold til digitalisering?</li><li>6. Hvor lenge pleier dere å være involvert i en digitaliseringsprosess?</li><li>7. Har dere opplevd et endret fokus innen digitalisering, gitt utviklingen de siste årene?</li></ol> |
| <b>Avslutning</b>                               | <ol style="list-style-type: none"><li>1. Er det noe mer du ønsker å legge til?</li><li>2. Takke for oss og for bidraget</li></ol>  |

## 9.4 Vedlegg 4: Intervjuguide #2 og #3

Før oppstart

- Presenterer gruppen og prosjektet og formålet ved intervjuet
- Går gjennom informasjonsskriv og innhenter skriftlig samtykke til deltagelse

| Tema  | Spørsmål  |
|---|---|
| <b>Intro</b>                                    | <ol style="list-style-type: none"> <li>1. Navn og alder</li> <li>2. Hva er din rolle/stilling?</li> <li>3. Kan du fortelle om din erfaring/ tidligere utdanning/bakgrunn?</li> <li>4. Hvor lenge har du vært ansatt/leder i bedriften?</li> </ol>   |
| <b>Spørsmål om data og teknologi</b>            | <ol style="list-style-type: none"> <li>1. Hva kreves av data for at bedrifter kan ta i bruk data-drevne prediksjoner?</li> <li>2. Hva er typiske problemer som oppstår mtp. datakvalitet?</li> <li>3. Hvordan hjelper dere bedrifter sikre god datakvalitet?</li> <li>4. Hvordan ser det typisk ut i en bedrift? Er ansvar sentralisert eller desentralisert?</li> <li>5. Hvordan henter virksomheter typisk inn data i dag? Er det vanlig for virksomheter å benytte seg av tredjepartsdata?</li> <li>6. Hvordan kan man best mulig prioritere eksterne datakilder? (f.eks. basert på relevante makrofaktorer, evt. bransje og næring)</li> </ol>                      |
| <b>Spørsmål om virksomhet</b>                   | <ol style="list-style-type: none"> <li>1. Hva burde vektlegges av bedrifter når det kommer til digitalisering?</li> <li>2. Hva er den vanligste grunnen til at bedrifter ønsker å digitaliseres?</li> <li>3. Opplever dere at det er et stort fokus fra ledelsen i virksomheter på å bli datadrevet?</li> <li>4. Er bedrifter gode på å være data-drevne?</li> <li>5. Opplever dere økt motstand mot digitalisering, gitt utviklingen de siste årene? Teknostress, etc.</li> </ol>  |
| <b>Spørsmål om fremgangsmåte i virksomheten</b> | <ol style="list-style-type: none"> <li>1. Ser dere en tydelig likhet mellom bedriftene som ønsker digitalisering?</li> <li>2. Hvor langt opplever dere at bedrifter er kommet i en digitaliseringsprosess, når de henvender seg til dere? Evt. Hvorfor?</li> <li>3. Ser du noen tydelige trender i hva bedriftene etterspør med tanke på teknologi?</li> <li>4. Hvor langt opplever dere at bedrifter ønsker å strekke seg ressursmessig, i forhold til digitalisering?</li> <li>5. Hvor lenge pleier dere å være involvert i en digitaliseringsprosess?</li> <li>6. Har dere opplevd et endret fokus innen digitalisering, gitt utviklingen de siste årene?</li> </ol> |
| <b>Avslutning</b>                               | <ol style="list-style-type: none"> <li>3. Er det noe mer du ønsker å legge til?</li> <li>4. Takke for oss og for bidraget</li> </ol>  |



## 9.5 Vedlegg 5: Intervjuguide #4

Før oppstart

- Presenterer gruppen og prosjektet og formålet ved intervjuet
- Går gjennom informasjonsskriv og innhenter skriftlig samtykke til deltagelse

| <b>Tema</b>                                      | <b>Spørsmål</b>   |
|--|---|
| <b>Intro</b>                                     | <ol style="list-style-type: none"><li>1. Navn og alder</li><li>2. Hva er din rolle/stilling?</li><li>3. Kan du fortelle om din erfaring/ tidligere utdanning/bakgrunn?</li><li>4. Hvor lenge har du vært ansatt/leder i bedriften?</li></ol>  |
| <b>Spørsmål om samarbeid mellom IT og BI</b>     | <ol style="list-style-type: none"><li>1. Hvordan samarbeider dere med dataanalytikere for å møte behovene deres mtp. arbeidet de utfører (BI-avdelingen)?</li><li>2. Hvilke utfordringer opplever dere ift. å møte disse behovene?</li></ol>  |
| <b>Spørsmål om tilgjengelig teknologi og opp</b> | <ol style="list-style-type: none"><li>3. Hva kreves for å hente inn og ta i bruk data om makroomgivelser? (f.eks ekstern data fra data brokers, data.norge)</li><li>4. Hvilke ressurser er tilgjengelige mtp. punkt 3?</li><li>5. Hvordan kan dere sørge for at IT-arkitekturen som understøtter avdelingen(e) ansvarlige for makroanalyser, er skalerbar og fleksibel nok til å håndtere fremtidige endringer iblant annet datakilder, analyseverktøy, teknologi, og brukerkrav?</li><li>6. Hva er mulige utfordringer knyttet til overnevnte punkt?</li></ol> |
| <b>Avslutning</b>                                | <ol style="list-style-type: none"><li>5. Er det noe mer du ønsker å legge til?</li><li>6. Takke for oss og for bidraget</li></ol>   |

## 9.6 Vedlegg 6: Meldeskjema for behandling av personopplysninger

08.01.2023, 23:04

Meldeskjema for behandling av personopplysninger



[Meldeskjema](#) / [Bachelorprosjekt for Digital Forretningsutvikling våren 2023](#) / Eksport

### Meldeskjema

#### Referansenummer

242976

#### Hvilke personopplysninger skal du behandle?

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidifikator
- Lydopptak av personer

#### Prosjektinformasjon

##### Prosjektittel

Bachelorprosjekt for Digital Forretningsutvikling våren 2023

##### Prosjektbeskrivelse

Formålet med prosjektet er å gjennomføre datainnsamlinger i det avsluttende bachelorprosjektet til studentene som følger studieprogrammet Bachelor i Digital Forretningsutvikling ved Institutt for datateknologi og informatikk ved NTNU våren 2023.

##### Begrunn hvorfor det er nødvendig å behandle personopplysningene

For å kunne studere anvendelse av IT med empiri fra en eller flere organisasjoner.

##### Ekstern finansiering

Ikke utfyllt

##### Type prosjekt

Studentprosjekt, bachelorstudium

##### Gjelder innmeldingen for flere studentprosjekter (felles vurdering)?

Ja

##### Oppgi antall studenter

46

#### Behandlingsansvar

##### Behandlingsansvarlig institusjon

Norges teknisk-naturvitenskapelige universitet / Fakultet for informasjonsteknologi og elektroteknikk (IE) / Institutt for datateknologi og informatikk

##### Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Leif Erik Opland, leif.e.opland@ntnu.no, tlf: 93254184

##### Skal behandlingsansvaret deles med andre institusjoner (felles behandlingsansvarlige)?

Nei

#### Utvalg 1

##### Beskriv utvalget

Utvalg fra en eller flere organisasjoner - må spesifiseres i hvert enkelt bachelorprosjekt.

##### Beskriv hvordan rekruttering eller trekking av utvalget skjer

Rekruttering skjer gjennom at en eller flere organisasjoner taes kontakt med, og det etterspørres en eller flere respondenter basert på et sett med utvalgsriterier for å sikre at respondentene er involvert i prosjekt- eller prosess man i prosjektet ønsker å studere. I videre kontakt med respondenter enten direkte eller via kontaktperson i bedrift, må studentene sørge for at både konfidensialitet og frivillighet ivaretaes.

**Alder**

18 - 67

**Personopplysninger for utvalg 1**

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidentifikator
- Lydopptak av personer

**Hvordan samler du inn data fra utvalg 1?****Personlig intervju****Vedlegg**[Mal intervjuguide v1.1.docx](#)**Grunnlag for å behandle alminnelige kategorier av personopplysninger**

Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

**Gruppeintervju****Vedlegg**[Mal intervjuguide v1.1.docx](#)**Grunnlag for å behandle alminnelige kategorier av personopplysninger**

Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

**Elektronisk spørreskjema****Vedlegg**[Mal sp\\_rreskjema.docx](#)**Grunnlag for å behandle alminnelige kategorier av personopplysninger**

Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

**Informasjon for utvalg 1****Informerer du utvalget om behandlingen av personopplysningene?**

Ja

**Hvordan?**

Skriftlig informasjon (papir eller elektronisk)

**Informasjonsskriv**[Informasjonsskriv-NSD v1.1.doc](#)**Tredjepersoner**

---

**Skal du behandle personopplysninger om tredjepersoner?**

Nei

**Dokumentasjon**

---

**Hvordan dokumenteres samtykkene?**

- Manuelt (papir)

**Hvordan kan samtykket trekkes tilbake?**

Ta kontakt med prosjektansvarlig å be om at samtykke trekkes tilbake.

#### Hvordan kan de registrerte få innsyn, rettet eller slettet personopplysninger om seg selv?

Gjennom å ta kontakt med prosjektansvarlig via opplysninger i informasjonsskriv.

#### Totalt antall registrerte i prosjektet

100-999

## Tillatelser

---

#### Skal du innhente følgende godkjenninger eller tillatelser for prosjektet?

Ikke utfyllt

## Behandling

---

#### Hvor behandles personopplysningene?

- Mobile enheter tilhørende behandlingsansvarlig institusjon
- Maskinvare tilhørende behandlingsansvarlig institusjon

#### Hvem behandler/har tilgang til personopplysningene?

- Student (studentprosjekt)
- Prosjektansvarlig

#### Tilgjengeliggjøres personopplysningene utenfor EU/EØS til en tredjestat eller internasjonal organisasjon?

Nei

## Sikkerhet

---

#### Oppbevares personopplysningene atskilt fra øvrige data (koblingsnøkkel)?

Ja

#### Hvilke tekniske og fysiske tiltak sikrer personopplysningene?

- Personopplysningene anonymiseres fortløpende

## Varighet

---

#### Prosjektperiode

09.01.2023 - 30.06.2023

#### Hva skjer med dataene ved prosjektslutt?

Data anonymiseres (sletter/omskriver personopplysningene)

#### Hvilke anonymiseringstiltak vil bli foretatt?

- Personidentifiserbare opplysninger fjernes, omskrives eller grovkategoriseres
- Lyd- eller bildeopptak slettes

#### Vil de registrerte kunne identifiseres (direkte eller indirekte) i oppgave/avhandling/øvrige publikasjoner fra prosjektet?

Nei

## Tilleggsopplysninger

---

## 9.7 Vedlegg 7: Vurdering av behandling av personopplysninger

08.01.2023, 22:59

Meldeskjema for behandling av personopplysninger



[Meldeskjema](#) / [Bachelorprosjekt for Digital Forretningsutvikling våren 2023](#) / Vurdering

### Vurdering av behandling av personopplysninger

**Referansenummer**  
242976

**Vurderingstype**  
Standard

**Dato**  
05.01.2023

**Prosjektittel**

Bachelorprosjekt for Digital Forretningsutvikling våren 2023

**Behandlingsansvarlig institusjon**

Norges teknisk-naturvitenskapelige universitet / Fakultet for informasjonsteknologi og elektroteknikk (IE) / Institutt for datateknologi og informatikk

**Prosjektansvarlig**

Leif Erik Opland

**Student**

ukjent

**Prosjektperiode**

09.01.2023 - 30.06.2023

**Kategorier personopplysninger**

Alminnelige

**Lovlig grunnlag**

Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene er lovlig så fremt den gjennomføres som oppgitt i meldeskjemaet. Det lovlige grunnlaget gjelder til 30.06.2023.

[Meldeskjema](#)

**Kommentar**

OM VURDERINGEN

Sikt har en avtale med institusjonen du forsker eller studerer ved. Denne avtalen innebærer at vi skal gi deg råd slik at behandlingen av personopplysninger i prosjektet ditt er lovlig etter personvernregelverket.

BAKGRUNN

Denne tilbakemeldingen gjelder en samlet vurdering av bacheloroppgaver. Denne vurderingen gjelder for studentoppgaver som følger retningslinjene som gis i denne tilbakemeldingen fra personverntjenester. Prosjekter som ikke følger de gitte retningslinjene må meldes inn på eget meldeskjema.

PERSONVERNTJENESTER SIN VURDERING

Prosjektansvarlig har ansvar for hvert enkelt prosjekt som omfattes av denne innmeldingen. Prosjektene skal gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg.

Dette betyr at studentene kan starte med datainnsamlingen.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektene vil behandle alminnelige kategorier av personopplysninger.

UTDYPENDE OM LOVLIG GRUNNLAG

Prosjektene vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektene 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.

KOMMENTARER TIL INFORMASJONSSKRIVET

Opplastet informasjonsskriv mangler noen punkter loven krever er med. Det må derfor legges til disse punktene i informasjonsskrivet for det gis videre til forskningsdeltakerne. Du trenger ikke å laste opp den oppdaterte versjonen i meldeskjemaet:

- Dato for prosjektslutt

- Oppdatert kontaktinformasjon for Sikt: Epost: [personverntjenester@sikt.no](mailto:personverntjenester@sikt.no)/telefon: 53 21 15 00

#### FØLG DIN INSTITUSJONS RETNINGSLINJER

Vi har vurdert at du har lovlig grunnlag til å behandle personopplysningene, men husk at det er institusjonen du er ansatt/student ved som avgjør hvilke databehandlere du kan bruke og hvordan du må lagre og sikre data i ditt prosjekt. Husk å bruke leverandører som din institusjon har avtale med (f.eks. ved skylagring, nettspørreskjema, videosamtale el. )

Personverntjenester 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).

#### MELD VESENTLIGE ENDRINGER

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til oss ved å oppdatere meldeskjemaet. Se våre nettsider om hvilke endringer du må melde: <https://sikt.no/melde-endringer-i-meldeskjema>

#### OPPFØLGING AV PROSJEKTET

Vi vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet. Prosjektansvarlig må forsikre seg om at studentene sletter rådata i forbindelse med innlevering/sensur av oppgavene. Det bør legges opp til at studentene bekrefter dette skriftlig til prosjektansvarlig når det er gjort, før prosjektansvarlig rapporterer om status for behandlingen av personopplysninger til personverntjenester.

Lykke til med prosjektet!

