

Business Intelligence for SMB

Helle Benjaminsen

Master i informatikk
Oppgaven levert: Juni 2011
Hovedveileder: Eric Monteiro, IDI

Forord

Denne masteroppgaven ble gjennomført ved Institutt for Datateknikk og Informasjonsvitenskap, ved Norges Teknisk- og Naturvitenskapelige Universitet, i perioden august 2010 til mai 2011.

Jeg ønsker å takke min veileder, Eric Monteiro, for gode og konstruktive tilbakemeldinger gjennom hele perioden. Jeg vil også takke Hilde Visthoff Drange, og Thomas Normann, for deres moralske støtte, og bidrag til oppgaven.

Takk til Acando, Conduct AS, Affecto, Steria, BearingPoint, Capgemini, Accenture, KnowIT, Bouvet, EDB Ergogroup, Deloitte, adEgo, Sherpa Consulting og Avanade, for bidrag til datainnsamlinger gjort i dette studiet.

Trondheim, mai 2011

Helle Sofie Thom Benjaminsen

Sammendrag

Denne oppgaven tar for seg små- og mellomstore bedrifters (SMB) behov for Business Intelligence (BI), og utfordringer for slike bedrifter. Det diskuteres hvordan disse bedriftene burde gå fram, for å lykkes med BI. Det gjøres også en vurdering på hvilke bedrifter som er egnet for å benytte seg av Open Source BI-programvare.

Innhold

1	Introduksjon	8
1.1	Business Intelligence	8
1.2	BI for SMB	9
1.3	Forskningsspørsmål	10
1.4	Oppgavens oppbygning	10
2	Business Intelligence	11
2.1	Business Intelligence' framvekst	11
2.2	Viktige elementer av BI	14
2.2.1	Fra rådata til kunnskap	14
2.2.2	Innhenting av rådata	16
2.2.3	Datastrukturene i BI	18
2.2.4	Uthenting av informasjonsdata	21
2.2.4.1	Rapportering	22
2.2.4.2	OLAP-analysering	22
2.2.4.3	Data mining	23
2.2.4.4	Dashboard	23
2.3	Motivasjoner for bruk av BI	24
2.3.1	Informasjonshåndtering	25
2.3.2	Informasjonsdeling	25
2.3.3	Operasjonell effektivitet	25
2.3.4	Metrikk-overvåkning	26
2.3.5	Strategi	27
2.3.6	Beslutningsstøtte	28
2.4	BI i praksis	29
2.4.1	Utfordringer	30
3	OSBI	32
3.1	Open Source	32

3.1.1	Open Source som utviklingsmetode	32
3.1.2	Andre typer programvare	33
3.2	Open Source Business Intelligence	34
3.2.1	Lisensieringsmodell	34
3.2.2	Arkitektur	35
3.2.3	OSBI-markedet	35
3.2.4	OSBI Vs. tradisjonell BI	36
3.2.5	Modenhet	37
3.2.5.1	Pentaho	40
3.3	Framtidsutsikter	40
4	2: SMB og BI	42
4.1	SMB	42
4.1.1	SMB i Norge	42
4.1.2	SMB vs. store bedrifter	43
4.2	SMB og BI	43
4.3	Utfordringer	44
4.3.1	Ansatte	44
4.3.2	Prosesser	47
4.3.3	Politikk og Prioritet	47
4.3.4	Oppsummering	48
5	Empirisk metode	50
5.0.5	Kvantitativ vs. Kvalitativ metode	50
5.1	Valgt tilnærming	52
5.2	Datainnsamling	53
5.2.1	Tilgang	53
5.2.2	Feltstudie	54
5.2.3	Spørreundersøkelse	55
5.2.3.1	Utforming av spørsmål	55
5.2.3.2	Pilottesting	56
5.2.3.3	Distribuering av undersøkelsen	56
5.2.4	Intervju	57
5.3	Dataanalyse	59
5.4	Refleksjon	59
5.4.1	Den hermeneutiske sirkelen	60
5.4.2	Kontekstualisering	60
5.4.3	Interaksjon mellom forsker og subjekter	61
5.4.4	Abstraksjon og generalisering	62
5.4.5	Flertolkninger	62

5.4.6	Misstanker	63
6	Feltstudie: Acandos BI-løsning	65
6.1	Bakgrunn	66
6.1.1	Acando	66
6.1.1.1	Steget mot BI	67
6.2	Pentaho BI Suite	68
6.2.0.2	Pentaho-versjonene	71
6.2.0.3	Eksempel på bruk av Pentaho BI Suite	71
6.3	Implementasjon	72
6.3.1	Kravspesifikasjon for rapporter	73
6.3.2	Database	74
6.3.2.1	Databasemodell	74
6.3.2.2	Kube	80
6.3.2.3	Aggregering av data	80
6.3.2.4	Databasetilkobling	81
6.3.3	Integrasjon	82
6.3.3.1	Active Directory	82
6.3.3.2	Sharepoint og Agresso	83
6.3.4	Rapportering	88
6.3.4.1	Fokusgruppe	88
6.3.4.2	Hente inn data	89
6.3.4.3	Rapport-design	90
6.3.5	Analysering	97
6.3.6	Aktørenes erfaringer og refleksjoner	98
6.3.6.1	Kundestyrt Prosjekt	98
6.3.6.2	Sommeransatte	99
6.3.6.3	Oppdragsgiver/veileder	100
6.3.6.4	Administrasjonen	101
6.4	Vurdering	101
7	BI-bransjen	103
7.1	Spørreundersøkelse	103
7.2	Intervjuer	105
7.3	Resultater	105
7.3.1	Respondentene	105
7.3.2	SMB Vs. Store Bedrifter	105
7.3.2.1	Erfaringsområder	106
7.3.2.2	Motivasjon om funksjonalitet	108
7.3.2.3	Motivasjon om informasjons-/dataflyt	110

7.3.2.4	Motivasjon om beslutningsstøtte og strategi .	111
7.3.3	OSBI vs. kommersiell BI	111
7.3.3.1	Kjennskap til de største OSBI-plattformene .	112
7.3.3.2	Antall brukere	113
7.3.3.3	Motivasjoner for bruk av BI	113
7.3.3.4	Valg av BI-plattform	115
7.3.3.5	Bruk av BI-funksjonalitet	117
7.3.3.6	Kilder brukt under utviklingen	118
7.3.3.7	Utfordringer for OSBI	120
7.3.3.8	OSBI i framtiden	121
7.3.3.9	OSBI for SMB	121
7.4	Funn	122
7.4.1	SMB's bruk av BI	122
7.4.2	OSBI	123
7.4.3	OSBI for SMB	123
7.4.4	OSBI i framtiden	124
8	Diskusjon	125
8.1	SMB's behov for BI	125
8.1.1	Automatisering av manuelle prosesser	125
8.1.2	Beslutningsstøtte og strategi	126
8.1.3	Analyse	127
8.1.4	Et helhetlig bilde	128
8.2	Utfordringer for SMB	128
8.2.1	Høy Risiko	128
8.2.2	Mangel på kunnskap om BI	129
8.2.3	Ansatte	130
8.2.4	Inkonsistente bedriftsmetrikker	131
8.3	Bruk av OSBI	131
8.3.1	Fordeler	131
8.3.2	Utfordringer	132
8.3.3	Modenhet	133
8.3.4	Åpen OSBI eller kommersiell OSBI?	134
8.3.5	OSBI for SMB	135
8.3.6	Framtidsutsikter	136
8.4	Anbefalt tilnærming til BI for SMB	137
8.4.1	Valg av programvare	137
8.4.2	Håndtering av risiko	138
8.4.3	Sikre kunnskap og kompetanse	138
8.4.4	Skape et realistisk utgangspunkt	139

8.4.5	Kommunikasjon	139
8.4.6	Tilrettelegge for vekst	140

9 Konklusjon 142

A Spørsmål fra spørreundersøkelse 145

A.1	Del 1 - Introduction	145
A.2	Del 2 - Background	145
A.3	Business Intelligence in practice	148
A.4	Del 4 - Open Source BI tools	154
A.5	Del 5 - Open Source BI Platforms	155
A.6	Avslutning	163

Kapittel 1

Introduksjon

1.1 Business Intelligence

Det er i dag vanlig for bedrifter å bruke flere forskjellige IT-systemer for håndtering av data. Dersom systemene ikke er integrerte med hverandre, er det vanskelig å se disse dataene i forhold til hverandre. Bedrifter som ønsker å generere rapporter fra de ulike systemene, vil da måtte bruke tid på manuell uthenting og sammenligning av data for hver eneste rapport som genereres; tid som ville vært spart dersom dataene var samlet i et integrert system. Hos noen bedrifter er systemene så mange, og dataene så uhåndterbare, at manuell uthenting ikke lar seg gjøre. Dette fører til at bedrifter besitter mye data, uten mulighet til å nå denne.

Mangelen på integrasjon, og uhåndterbar data, blir et problem dersom man ønsker å gjøre analyser på datagrunnlaget. Bedrifter har helt siden den digitale dataens oppstandelse sett et behov for å bruke tilgjengelig data for å ta bedriftsrelaterte avgjørelser. Systemer for å gjøre slik integrasjon og analyse mulig, har eksistert i flere tiår. Tidligere snakket man om “IT-basert ledelse-støtte”, “Ledelsestøttesystemer” og “Ledelse-informasjonssystemer” og disse hjalp bedriftsledere å ta avgjørelser [32]. I dag går alle disse termene under en felles term; *Business Intelligence* (heretter omtalt som BI). BI beskrives som en bedriftsprosess hvor man *samler*, *analyserer* og *presenterer* data, for å oppnå et gitt formål [71].

Ved å benytte BI kan bedrifter få presentert informasjon på en ny, og mer oversiktlig måte. Ulike systemer integreres, og dataen fra disse samles i et såkaldt *datavarehus*. Dataen som lagres her, blir lagret rundt ulike subjektområder, som skal gjøre det enklere å analysere data rundt gitte KPI-er/bedriftsmålinger (eks. omsetning, kundevekst, markedsvekst ol.). I

datavarehuset samles data over tid, slik at man får et historisk datagrunnlag med informasjonsdata. På denne måten blir det mulig å analysere dataens oppførsel over tid, og se etter mønster i datagrunnlaget. I tillegg til analysering, er *rapportering* et av de største funksjonelle områdene innen BI. Automatisk rapportering sparer mye tid for de ansatte, og gjør informasjonsdeling lettere og mer effektivt.

Et eksempel på bruk av BI er et telekom-selskap som ønsker å føre en strategi basert på bruksmønster, samtidig som de vil gjøre manuell rapportering mer effektivt. De benytter da BI for å integrere kundesystem, faktureringsystem og økonomisystem til et samlet datavarehus. Når dataen ligger i datavarehuset kan de benytte BI-funksjonalitet for å lage ferdige rapporter som skal genereres automatisk, og sendes ut til mottaker. Et eksempel på en slik rapport er faktura til kunde, som lages automatisk hver måned. I tillegg bruker bedriften BI-funksjonalitet for å gjøre analyser på bruksmønstrene til kundene. Ved å analysere ulike bruksmønster kan de identifisere og klassifisere ulike brukergrupper, og lage nye typer abonnementer som skal passe disse gruppene. På denne måten ønsker de å oppnå at kunder blir mer fornøyde med abonnementet sitt, og dermed blir værende som kunde.

1.2 BI for SMB

BI var i starten hovedsakelig benyttet av store bedrifter. Prosessen med å integrere systemer og designe et datavarehus er svært omfattende, i tillegg til at BI-programvaren har vært dyr. Dette har bidratt til å ekskludere små- og mellomstore bedrifter (heretter omtalt som SMB) fra BI-markedet. Mindre bedrifter har lavere omsetning, og vil derfor få en høyere risiko for investering i nye teknologier, som BI. Dette har ført til at SMB-er som har sett et behov for BI, ikke har hatt mulighet til å investere i BI, grunnet den høye prisen.

De siste årene har BI-programvaren sunket betraktelig i pris. Det har i tillegg oppstått flere Open Source BI-alternativer på markedet, noe som har ført til at flere SMB-er tørr å ta steget mot BI. Men selv om programvaren har blitt billigere, har fortsatt SMB-er en risiko knyttet til investering av BI, da en mislykket BI-implementasjon kan få mye å si for bedriften økonomisk. Det er derfor interessant å se på hvilke utfordringer SMB-er står overfor, og hvordan disse burde håndteres, for at SMB-er skal lykkes med BI. Denne oppgaven vil forsøke å besvare disse spørsmålene, samtidig som det gjøres en vurdering på hvorvidt et Open Source BI-alternativ er modent nok for bruk av SMB-er.

1.3 Forskningsspørsmål

Forskingsspørsmål i denne oppgaven er 1) Hvilke behov SMB har for BI, 2) Hvilke utfordringer SMB-er som ønsker å implementere BI står overfor, 3) Hvilke bedrifter som kan benytte OSBI, og 4) Hvordan SMB-er som ønsker å implementere BI burde gå fram for å lykkes.

1.4 Oppgavens oppbygning

Denne oppgaven starter med et kapittel som beskriver Business Intelligence som teknologi, motivasjoner for bruk, og bruk av Business Intelligence i praksis. Kapittel 3 tar for seg SMB som markedssegment på BI-markedet, og utfordringer SMB-er står overfor. Kap 4 omhandler Open Source Business Intelligence, med en sammenligning av OSBI og tradisjonell BI, og en vurdering av OSBI's modenhet.

Kap 5 beskriver de empiriske metodene som ble brukt i datainnsamlingen i dette studiet. Kap 6 tar for seg et feltstudie som ble gjennomført på en SMB som implementerte en OSBI-løsning, og kap 7 beskriver innsamlet data fra BI-markedet, fra en spørreundersøkelse og intervjuer.

Kap 8 diskuterer funnene i kap 6 og 7, og forsøker å besvare forskningsspørsmålene i denne oppgaven. Til slutt følger en kortfattet konklusjon fra dette studiet.

Kapittel 2

Business Intelligence

Business Intelligence har de siste 20 årene blitt stadig mer populært. Det har ved hjelp av BI blitt mulig å se data på nye måter, ved å få ny informasjon om allerede eksisterende data. Dette gir ledelsen i bedrifter mulighet til å få hjelp i forretningsstrategiske avgjørelser. Man har gått fra å bruke data til å få svar på spørsmål som “*hva skjer nå?*” til å ved hjelp av BI også få svar på spørsmål som “*hvorfor har dette skjedd?*” og “*hva vil skje i framtiden?*” [71].

2.1 Business Intelligence’ framvekst

BI’s opprinnelse kan spores helt tilbake til da applikasjoner kjørte på sekvensiell teknologi, som for eksempel magnetbånd [42]. På den tiden måtte hele filen aksesserer, selv om det kun var behov for bare en liten del av fila. Man snakket da om “master-filer” som holdt data som ble aksessert fra flere applikasjoner. Ofte ble båndene slitt, og hele filer gikk tapt. Dette førte til et økende behov for en ny type fillagring, for sikrere lagring av data. Etter hvert ble magnetbåndene byttet ut med disk, og dette førte til at data kunne aksesserer direkte, noe som økte effektiviteten i dataauthenting betraktelig. Prosessorer ble kraftigere, og hastighet og pris på disse sank raskt. Master-filene ble byttet ut med databaser; sentraliserte samlinger med data. Disse ble lagret på disk, tilgjengelig for prosessering av de applikasjoner som trengte de. Transaksjoner kunne nå prosesseres direkte mot den sentraliserte databasen, og resultatet ble sendt over internett. Muligheten for denne online-transaksjonsprosesseringen førte til at online-applikasjoner oppstod, og disse ble knyttet sammen med de sentrale online-databasene, som ble brukt for lagring. Denne nye typen applikasjoner hadde høyt fokus på tilgjen-

gelighet og god responstid. Mange bedrifter gjorde online-applikasjonene sentrale for sin virksomhet, og de ble en essensiell del av interaksjonen mot kunder.

Utviklingen av online-transaksjonsapplikasjoner førte til flere utfordringer:

- Et stort antall applikasjoner
- Mangel på en felles forståelse for hvordan en applikasjon skulle se ut.
- Lite/ingen dokumentasjon
- Stadig nye krav
- Inhabilitet til å endre applikasjoner etter fullført utvikling
- Sårbare systemer
- Sensitiv responstid
- Mangelfull/fraværende integrasjon mellom applikasjoner

Et resultat av den mangelfulle integrasjonen mellom applikasjonene [42], var at det ikke fantes en felles forståelse av hva slags data som fantes, og hvordan denne kunne/burde brukes. Denne begrepsforståelsen omhandlet begreper som for eksempel *kunde*, *produkt*, *transaksjon*, *leverandør* og *frakt*. Siden applikasjonene var svært vanskelig å endre, var det nærmest umulig å gjøre omtolkninger av disse begrepene i applikasjonene. Som resultat av dette, hadde selv de største og mest sofistikerte bedriftene problemer med å finne helt grunnleggende informasjon, som hvem kundene deres faktisk var. Man hadde store mengder data, men ingen mulighet til å hente ut denne. Man fikk dermed ikke mulighet til å måle metrikker som for eksempel hvordan produkter som solgte best, hvordan kunder som var de beste kundene, og hvor store inntekter man hadde hatt per kvartal.

Første reaksjon på problemet rundt mangel på forretningsdata [42], var å lage såkalte ekstraktfiler. Disse ble laget av en database fra en applikasjon, og sendt videre til neste applikasjon. På den måten var det tilsynelatende mulig å dele data imellom applikasjoner, og å lage forretningsdata. Ekstraktfiler ble raskt populære, men de skapte flere problemer enn løsninger på integrasjonsproblemet; for hver nye ekstraktfil som ble laget, ble spindelvevet av ekstraktfiler utvidet, noe som ga konsekvenser:

- Mangel på dataintegritet: Et dataelement hadde flere versjoner/verdier på ulike steder, og ingen visste hva som var rett verdi

- Dataredundans: Redundansen av dataen var enorm, ettersom den ble kopiert på nytt for hver gang en ekstraktfil ble sendt videre til neste applikasjon
- Eldende data: Mens dataen ble sendt gjennom systemet ble dataen utdatert. En verdi kunne endres 5-6 ganger i løpet av en dag, og prosessene for dataauthenting var ute av stand til å holde dataen oppdatert ved endring
- Flere siloer med data: Hver silo hadde et eget domene, uten koordinasjon eller integrasjon med andre siloer. Dette førte til at beslutninger tatt i ett domene ofte var i strid med andre deler av organisasjonen.
- Systemet frøs: Ekstraktfilene førte til at allerede kollapsende systemer frøs. Applikasjonene var allerede vanskelige å endre, og ved å “pakke” de inn i ekstraktfiler ble endringer helt umulige

Disse online-applikasjonene var spesialiserte på å få tak i nåværende data, som for eksempel “hvor er forsendingen nå?”, eller “hvor stor saldo er på konto nå?”. Historisk data ble ofte slettet, ettersom den “tettet igjen årene” til den ferskeste dataen [42], og svekket effektiviteten til online-prosesseringen. Mangelen på historisk data førte til at man ikke hadde mulighet til å gjøre prediktive dataanalyser. Mange organisasjoner ønsket å forstå kundebasen sin, ettersom kunder har vaner som kan gjenkjennes ved å se på historiske data i form av mønster. Dette var ikke mulig da det i liten grad fantes integrert og historisk data i systemet, og selv om denne dataen ble funnet kunne man ikke vite om den var pålitelig. Organisasjoner ble stadig mer frustrerte, ettersom de ofte visste at dataen fantes et sted i systemet, men de greide ikke å nå den. Det ble tydelig at man trengte en ny måte å håndtere data på, og denne måtte støtte lagring og analysing av historisk data. Man definerte på denne tiden to typer data; *operasjonell data* og *informasjonsdata*.

Operasjonell data er den typen data som brukes i transaksjonsprosesser¹, slik som beskrevet tidligere. Denne typen data bidrar til å støtte dagligdagse operasjoner, ved å ha oppdatert applikasjonsorientert data, med god tilgjengelighet. *Informasjonsdata* er mer summerende, og mindre detaljert enn operasjonell data. Den er hovedsakelig integrert, historisk data, og brukes som et grunnlag for analyser for forretningsavgjørelser. Det er ikke mulig å oppdatere informasjonsdataoppføringer, da man kun kan legge til nye data. Hver dataoppføring har et tidsstempel, som angir et tidspunkt eller en periode for når dataen stemte. Informasjonsdataen er designet rundt et

¹Prosesseringer som gjøres i tradisjonelle ERP-systemer.

subjektområde, i motsetning til operasjonell data som er designet rundt et funksjonelt område. Dette vil beskrives nærmere i seksjon 1.2.1.

Den nye typen data, informasjonsdata, var starten på det vi i dag kjenner som datavarehus. På den tiden (70-tallet) gikk det under navnet beslutningsstøttesystem. Termen Business Intelligence ble først brukt på tidlig 90-tallet [48], og ble ofte brukt som et synonym til datavarehus og beslutningsstøttesystem. I dag er det en felles forståelse av at Business Intelligence er et strukturert informasjons-aksesslag som finnes imellom datavarehuset og sluttbrukeren. Likevel er det som oftest også snakk om datavarehus når man snakker om BI, ettersom en BI-løsning bygges på toppen av et datavarehus.

2.2 Viktige elementer av BI

2.2.1 Fra rådata til kunnskap

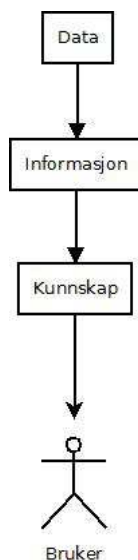
Business Intelligence beskrives som en bedriftsprosess hvor man *samler*, *analyserer* og *presenterer* data, for å oppnå et gitt formål [71]. Definisjonen av BI er svært abstrakt, noe som er et resultat av at BI kan være mye forskjellig, avhengig av hva man ønsker å oppnå. Generelt kan man si at BI gjør det mulig å gjøre data om til informasjon, og deretter gjøre informasjonen om til kunnskap.

I følge Davenport og Prusak[15] blir data omgjort til informasjon ved hjelp av følgende metoder:

- Kontekstualisering: Meningen med dataen blir kjent
- Kategorisering
- Kalkulering: Dataen analyseres
- Rettelse: Feil fjernes fra dataen
- Summering

Data som skal gjøres om til informasjon kommer typisk fra flere ulike kilder. Man siler ut den dataen man ønsker fra hver kilde, plasserer den i en kontekst, kategoriserer den og kvalitetsikrer den mot feil. Deretter legges informasjonsdataen inn i informasjonsbasen; *datavarehuset*. Dataen i datavarehuset vil ikke oppdateres med “ferskere” data, men blir samlet over lang tid. Prosessen, hvor data fra kildesystemene transformeres og overføres som informasjonsdata til datavarehuset, kalles *ETL*². Når informasjonen ligger

²Extract, transform, load.

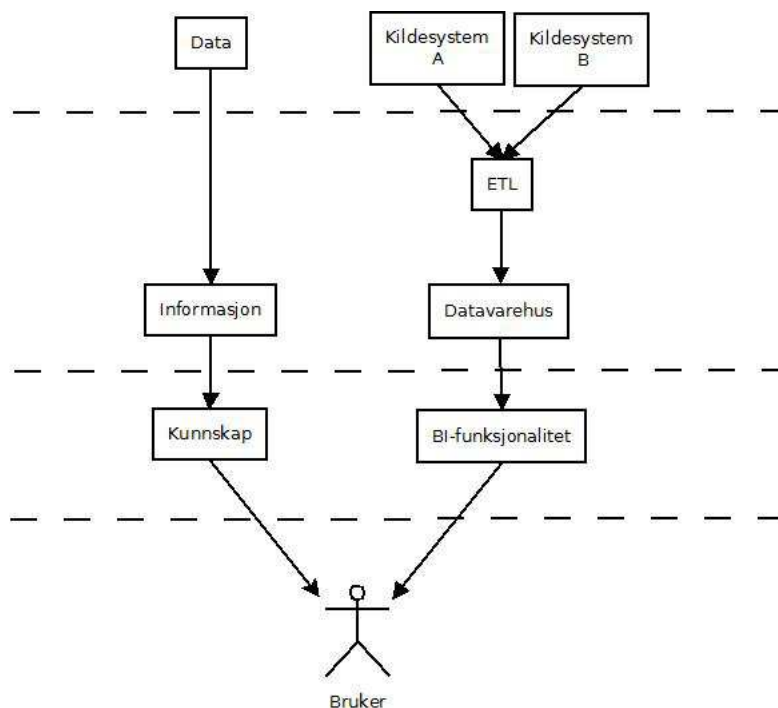


Figur 2.1: Figuren illustrerer hvordan data først må gjøres om til informasjon, før den kan gjøres om til kunnskap.

i datavarehuset, kan den trekkes ut og gjøres om til kunnskap. Davenport og Prusak [15] mener at informasjon gjøres om til kunnskap ved hjelp av følgende metoder:

- Sammenligning
- Finne sammenhenger
- Finne konsekvenser
- Samtale

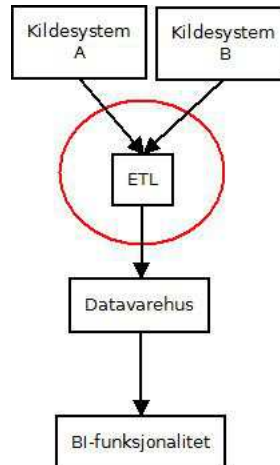
BI-verktøy er designet for å gjøre det så enkelt som mulig for brukeren å analysere informasjonsdataen i datavarehuset, ved å se på sammenhenger og mønster i den historiske informasjonsdataen. Man oppnår da kunnskap ved å benytte BI-verktøyene. BI gjør det også mulig å presentere informasjonsdataen, ved hjelp av et bredt utvalg av verktøy.



Figur 2.2: Figuren viser hvordan data omgjøres til kunnskap ved hjelp av BI.

2.2.2 Innhenting av rådata

Data som skal benyttes kan komme fra flere ulike datalager/kilder. Dette kan være ERP-systemer, databaser, kommaseparerte filer, excel filer, osv [32]. ETL er forkortelsen for de tre primærstegene som samler data fra kildesystemene og laster disse inn i datavarehuset; *Ekstrahering, transformering* og *lasting*. Det finnes mange gode verktøy for å utføre ETL, og de fleste BI-leverandører leverer spesialprogramvare for å utføre disse tre stegene [42].



Figur 2.3: ETL- steget henter data fra kildesystem og integrerer disse inn i et sentralisert datavarehus.

Det første steget, *ekstrahering*, består av uthenting av data fra kilde. Dette gjøres ved å sette opp koblinger mot kilder som holder dataen, og lagre dataen midlertidig i en såkalt ekstraksjon, som fungerer som et hurtigminne/cache [32].

Midtsteget i integrasjonsprosessen *transformerer* innhentet data fra opprinnelig til ønsket format [32]. En av de største tekniske utfordringene ved denne transformasjonen er å skape et felles vokabular for de ulike kildene [19]. Et eksempel er to systemer, et kundesystem og et faktureringssystem, som inneholder kunde- og faktureringsdata. I kundesystemet er kunde-ID noe annet enn i faktureringssystemet. Det gjennomføres derfor en sjekk, ved å analysere navn-strengen i faktureringssystemet med navn-strengen i kundedatabasen. Man får da et felles vokabular for kunde-ID, og kan videre integrere feltene som omhandler kunde. Det settes også et tidsstempel på all data, for at det skal være mulig å se hvordan dataen endrer seg over tid.

Oppsummert innebærer transformeringen følgende aktiviteter [42]:

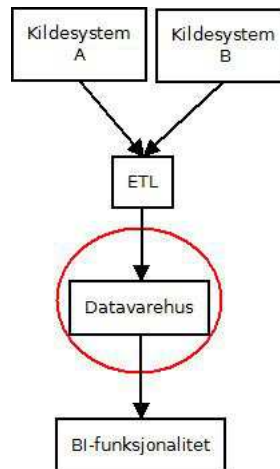
- Benytte filter for å sile ut uønsket og korrupt data
- Utføre syntaktisk og operasjonell synkronisering/fletting av data fra ulike kilder
- Konvertering av nøkler
- Sette tidsstempel på data

- Reformatering av dataen

Etter fullført transformering *lastes* ferdigtransformert data inn i datavarehuset.

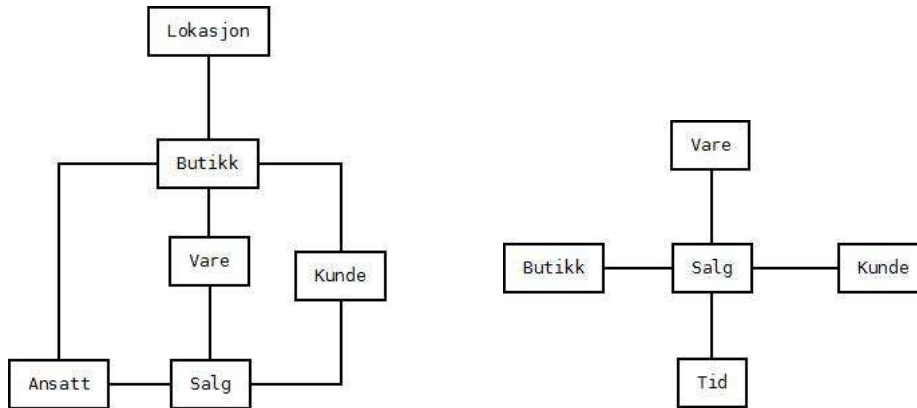
2.2.3 Datastrukturene i BI

BI-applikasjoner må være i stand til å håndtere store mengder data raskt. Effektivitet i datastrukturene som lagrer dataen i datavarehuset er derfor svært viktige.



Figur 2.4: Informasjonsdataen som lagres i datavarehuset lagres i spesialtilpassede datastrukturer.

Et datavarehus lagrer data historisk, over tid. Det betyr at datavarehuset bare vil vokse, siden data aldri slettes. Dette er ikke vanlig i tradisjonelle operasjonelle systemer, som prosesserer dag-til-dag operasjoner, og holder midlertidig data. Operasjonelle systemer har fokus på å bruke liten lagringsplass, for å få rask data-innsetting og -uthenting for et stort antall brukere. Dette er altså det motsatte av fokus for datavarehus, hvor lagringsplass ikke er et problem, og brukere ikke har mulighet til å sette inn ny data (all ny data hentes ifra andre systemer/kilder). Operasjonelle systemer bruker ER-modellering og normalisering for database-design, mens BI-systemer bruker en dimensjonal databasemodell, som vist nedenfor.



Figur 2.5: Modellen t.v. viser en typisk ER-database for butikkssalg, mens modellen t.h. viser butikkssalg lagret i en dimensjonal database.

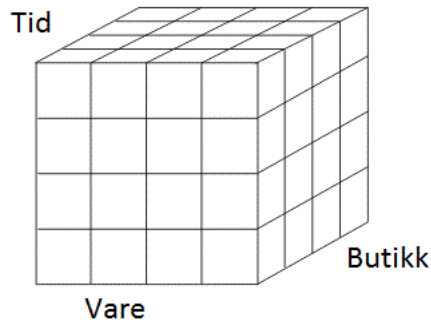
Ved å normalisere dataen i ER-modellen blir så lite plass som mulig brukt, ved at det er få kopier av samme data. Den dimensjonale modellen tar lite hensyn til replikering av data, noe som gjør database-strukturen lettere å lese. Dette er nyttig for brukere som ønsker å skrive sine egne spørringer mot databasen, og dermed må skjønne datastrukturen.

Den dimensjonale modellen lagrer data i et *stjerne-skjema*³, med en *fakta-tabell* i midten, og *dimensjonstabeller* rundt. Fakta-tabellen holder typisk en måling for en aktivitet, mens dimensjonene beskriver aktiviteten. I eksempelet over er fakta-tabellen og aktiviteten “Salg”, og tabellene rundt beskriver tidspunkt og butikk hvor salget fant sted, samt hvordan vare som ble solgt. Måling-feltet i faktatabellen er summen⁴ av salget. Tidsdimensjonen inkluderes alltid i et stjerneskjema, for at det skal være mulig å se hvordan dataen endres over tid.

En annen datastruktur som brukes i de fleste BI-løsninger er *OLAP-kuber*. OLAP-kuber er multidimensjonale datastrukturer som, til tross for navnet, kan ha mer enn 3 dimensjoner. Disse kubene bygges på toppen av stjerneskjema-modellen. I visualiseringen av kubene nedenfor er det for enkelthets skyld kun tatt med 3 av de 4 dimensjonene til salgsdatabasen.

³Navnet kommer av at skjemaet ser ut som en stjerne med et midtpunkt og tagger rundt.

⁴Salget er i dette tilfellet *subjektområdet* til det som skal lagres, og alle analyser som gjøres på dette skjemaet vil skje ved å se på summen av salg på ulike nivå.



Tabell 2.1: Kube for produktsalgs-databasen.

Tidsdimensjonen i stjerneskjemaet benyttes i kuben for å opprette et *tids-hierarki*. Det blir på denne måten enkelt å “zoome inn” på dataen på gitte tidsintervaller (dager, uker, måneder, år). MDX er språket som benyttes for å gjøre spørringer mot OLAP-kuber, på samme måte som SQL brukes for å gjøre spørringer mot relasjonsdatabaser [73].

Når kuben genereres vil summert data forhåndskalkuleres for de ulike dimensjonene, slik at spørringer går raskere. Slike forhåndskalkuleringer kalles *aggregater*.

Ved å kombinere stjerneskjema og kube vil det være enkelt å skrive spørringer, samtidig som det går svært raskt.

Eksempler på spørringer mot salgs-databasen, hvor både A, B, og C er ulike aggregater lagret i kuben:

- Sum av salg av vare A som fant sted i år B
- Sum av salg av alle varer som ble kjøpt av kunde A i måned B
- Sum av salg av vare A som ble solgt i alle butikker på dag B
- Sum av salg av vare A i kvartal B, sammenlignet med kvartal C

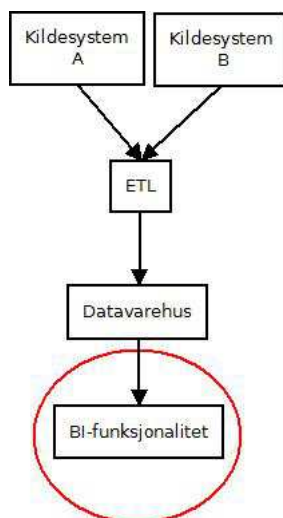
Det blir i disse eksemplene tydelig at det er *summen* av hvert salg som er av interesse, da det er nettopp denne som er i sentrum av stjerne-skjema, og som med ulike summeringer lagres i kuben.

Data som ligger på laveste nivå i databasen kalles *granular data*, og denne ligger som regel ikke i aggregater. Ved å ha granular data som basis i datavarehuset blir det mulig å se på dataen i fra flere synsvinkler. Et eksempel er subjektområdet *salg*, som vi nettopp så på, hvor flere brukergrupper ønsker å se på salgsdataen på ulike måter. Markedsavdelingen, finansavdelingen og salgsavdelingen er alle interesserte i ulike aspekter av salgsdataen,

men de bruker likevel samme kilde data. Vi har nettopp sett på et eksempel for markedsavdelingen, mens stjerneskjema og kube for de andre brukergruppene ville sett annerledes ut, men med mye av den samme dataen. En annen fordel med granular data, er at denne kan lagres i datavarehuset over tid, selv om det ikke er bruk for den for øyeblikket. Dermed ligger den klar for framtidig bruk, dersom det skulle oppstå nye bedriftskrav (og dermed også krav til informasjon). Da vil man slippe å oppsøke det operasjonelle miljøet for å få ut denne dataen, og da risikere at denne er slettet. Bedriften vil på denne måten få en proaktiv posisjon, i stede for en reaktiv posisjon, i forhold til nye krav om informasjon.

2.2.4 Uthenting av informasjonsdata

Datavarehuset er en veldig fleksibel informasjonskilde [56]. Det er mulig å skrive spørringer direkte mot datavarehuset, men BI-leverandører tilbyr mange ulike verktøy for å trekke ut og presentere denne informasjonsdataen.



Figur 2.6: Når datavarehuset er fylt med informasjonsdata, er det klart for å trekke denne ut ved hjelp av BI-funksjonalitet.

BI-funksjonaliteten er som regel lik hos de ulike BI-leverandørene, og består av følgende hovedfunksjonaliteter:

- Rapportering

- OLAP-analysering
- Data mining
- Dashboarding

2.2.4.1 Rapportering

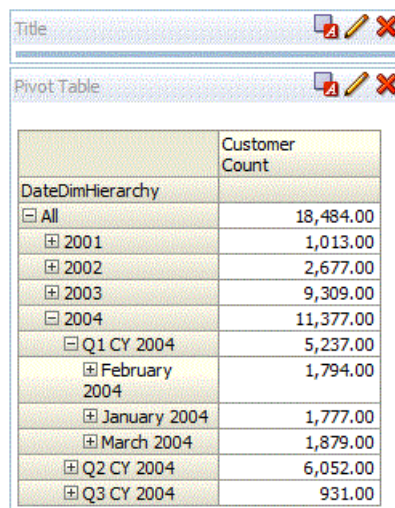
Rapporteringsverktøy var det første datavarehus-verktøyet, og selve årsaken til at datavarehus ble så populært i utgangspunktet [42]. Slike verktøy brukes for å summere statusene til bedriften, ved å generere ulike rapporter. Disse kan genereres ved satte intervaller, eller på kommando fra bruker. Eksempler på rapporter er rapport for omsetning, rapport for antall timer jobbet per ansatt, og rapport for antall varer igjen på lageret. Et mer avansert eksempel er en rapport som genereres dersom en hendelse inntreffer. Det kan være en rapport som sendes ut til arbeidsgiver dersom en ansatt har jobbet mer overtid enn hva han har lov til. En slik rapport baseres på en analyse, og må kjøres ved gitte intervaller, for å undersøke om hendelsen har inntruffet. Disse rapportene lages typisk av teknisk personell igjennom rapporteringsverktøy. Slike verktøy har drag-and drop-grensesnitt for å legge til elementer i rapportern. Det gjøres en oppkobling mot datavarehuset, og spørringer for å hente ut data gjøres igjennom dette verktøyet. Dataen som hentes ut igjennom spørringene brukes som parametre i rapportene. På denne måten blir dataen i rapportene dynamiske, og vil endres når data-grunnlaget endres.

En annen type rapportering som støttes av de fleste BI-plattformer er *ad-hoc rapporter*, som skal gjøre det enkelt for ikke-tekniske brukere å lage rapporter. Dette gjøres ved at teknisk personell først lager en halv-ferdig rapport i rapportdesigneren, og deretter gjør denne tilgjengelig i et grensesnitt for brukeren, som regel igjennom et web-grensesnitt. Da kan den ikke-tekniske brukeren velge mellom et utvalg av datafelt og operasjoner i ad-hoc rapporten.

2.2.4.2 OLAP-analysering

Analyseverktøyene som tilbys av BI-leverandørene er basert på OLAP⁵, som gjør det mulig å analysere store mengder data, svært raskt. Styrken til OLAP ligger i den multidimensjonale kube-datastrukturen, som beskrevet tidligere. Ved å benytte kuber blir det mulig å analysere dataen grafisk, ved å gjøre en såkalt *drill-down* i datagrunnlaget, for å finne det detalj-nivået man ønsker.

⁵Online Analytical Processing.



The image shows a software window titled "Pivot Table" with a drill-down tool. The table displays the following data:

DateDimHierarchy	Customer Count
All	18,484.00
2001	1,013.00
2002	2,677.00
2003	9,309.00
2004	11,377.00
Q1 CY 2004	5,237.00
February 2004	1,794.00
January 2004	1,777.00
March 2004	1,879.00
Q2 CY 2004	6,052.00
Q3 CY 2004	931.00

Figur 2.7: Et eksempel på et drill-down verktøy, som i dette tilfellet viser antall nye kunder i ulike perioder.

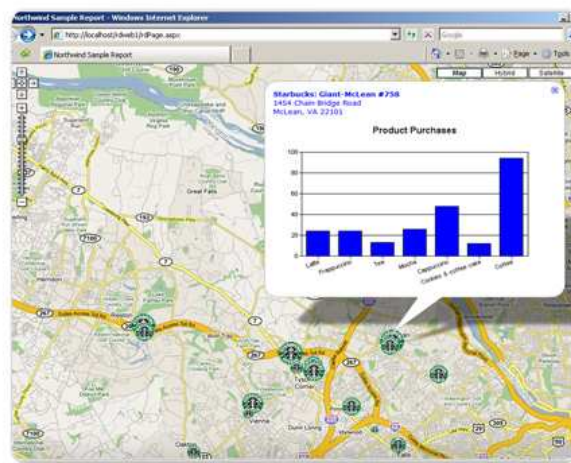
2.2.4.3 Data mining

Data mining er en term som brukes om mønstergjenkjenning i datagrunnlaget. Ved å koble sammen ulike datakilder, og se på disse i forhold til hverandre, kan man oppdage mønster som ikke var opplagte ved første øyekast. Dette kan gi kunnskap om historisk data, som igjen kan gi indikasjoner på hvordan framtidig data vil se ut. Denne typen analyse skiller seg ut ved at man ikke vet på forhånd hva man ser etter. Data mining kan brukes for å optimalisere og forbedre det som analyseres. Et eksempel er en butikk-kjede som ser på en vares lokasjon i forhold til hvor mye det selges av den gitte varen. Ved å plassere varen på ulike steder i butikken, kan man over tid finne ut ved hvilken plassering varen selger best.

2.2.4.4 Dashboard

Dashboard-funksjonalitet gjør det mulig å lage visuelle rapporter som kan endres ved input fra bruker. Enkle eksempler på dashboards er kake- og stolpediagrammer, som gir mer informasjon når brukeren holder musepila over de ulike delene av diagrammet. Et mer avansert eksempel er et kart, hvor brukeren kan trykke på et land for å få opp mer informasjon om det valgte landet.

Dashboards kan også lages uten at de nødvendigvis tar input fra bruk-



Figur 2.8: Et eksempel på et dashboard som viser et kart med data for ulike lokasjoner.

eren. Nedenfor vises en måler for salg i forhold til budsjett.



Figur 2.9: Eksempel på dashboard i form av en dynamisk måler.

2.3 Motivasjoner for bruk av BI

- Informasjonshåndtering: Samle data fra flere kilder, til en felles informasjonsbase. Dette gjøres for å skape et godt analytisk grunnlag
- Informasjonsdeling: Et felles informasjonslager gjør det lettere å spre kvalitetssikret informasjon gjennom rapport-distribuering

- Operasjonell effektivitet: Automatisering av manuelle prosesser, og tilrettelegging for en selvbetjeningsmodell for brukere
- Metrikk-overvåkning: Overvåkning av KPI-er, og ansattes prestasjoner.
- Strategi: Føre en BI-basert strategi
- Beslutningsstøtte: Basere bedriftsavgjørelser på BI-analyser

2.3.1 Informasjonshåndtering

En av de viktigste motivasjonene for bruk av BI, er å få samlet og kvalitetssikret data fra ulike kilder. Mange bedrifter sliter med at data som skal beskrive det samme subjektområdet, er forskjellig hos ulike kilder. Ved å ha kun ett sentralt datalager, med kvalitetssikret data, kan man være sikker på at det ikke finnes andre versjoner av den samme dataen; det finnes bare *én sannhet* om dataen. Man strukturerer dataen som informasjon, ved å lagre de i stjerneskjema-strukturer, som nevnt i Seksjon 1.2.3. Dette gjør det lettere å nå dataen/informasjonen, samtidig som man kan være sikker på at den er korrekt. Ved hjelp av datavarehus gjør man ustrukturert data om til informasjon. Det blir da videre mulig å gjøre informasjonen om til kunnskap, ved hjelp av ulike BI-verktøy.

2.3.2 Informasjonsdeling

Mulighet til å dele informasjon med kunder, ansatte, partnere o.l. er en vanlig motivasjon for bruk av BI. Med ett felles datalager blir det tryggere og enklere å dele data og informasjon, ved hjelp av rapporter. Deling av rapporter kan skje på mange ulike måter; de kan for eksempel aksesseres gjennom web-browseren, vises som dashboards på skrivebordet, eller bli sendt som pdf-fil på e-post. Rapporteringsfunksjonaliteten gjør det mulig å spre informasjon som tidligere kun var tilgjengelig for ansatte som jobbet spesifikt med rapportering, og/eller hadde teknisk kompetanse.

2.3.3 Operasjonell effektivitet

Ved å automatisere generering av rapporter, som tidligere ble laget manuelt, oppnår man operasjonell effektivitet, ettersom rapport-ansvarlige da kun vil trenge å lage rapporter én gang [27]. Ved at rapportene har dynamiske parametre som automatisk hentes inn fra det kvalitetssikrede datavarehuset, vil man unngå at informasjonen i rapportene blir feilaktig, samtidig som

rapportene endres dersom datagrunnlaget endres. Det blir da ikke lenger en risiko at rapportansvarlig registrerer feil data, som kan resultere i feil analyseresultat, som igjen kan medføre at bedriften tar forretningsbeslutninger basert på feil informasjon [27].

Et eksempel på autogenerering av en rapport, er en varebestillingsrapport, som sendes til bestillingsansvarlig i en matbutikk ved et gitt intervall. Bestillingsansvarlig får da en ferdig rapport som viser hvor mange varer av en gitt sort som må bestilles. Denne rapporten kan baseres på hvor mange varer som er igjen i butikk, samt ved å se på mønster for hvor mange av en gitt vare som ble solgt den tilsvarende perioden året før. Butikken vil for eksempel selge flere engangsgriller i august enn i september. Ved å autogenerere en slik rapport, får bestillingsansvarlig hjelp til å vite når en vare på bestilles, og slipper å bruke tid på å sjekke, og vurdere dette manuelt.

Halvferdige ad-hoc rapporter er, som nevnt i Seksjon 1.2.4.5, en annen type rapport. Ved å gjøre slike rapporter tilgjengelige for brukere, oppnår man en selvbetjeningsmodell for brukerne. De slipper da å være avhengige av teknisk personell for å få ut rapporter, og kan selv enkelt konfigurere hva som skal inkluderes i rapporten. Dette er tidsbesparende, både for bruker og for rapport-ansvarlige [27], og fører dermed til operasjonell effektivitet.

2.3.4 Metrikk-overvåking

Ved å ha mulighet til å automatisere prosesser, skapes mange muligheter. Dersom en bedrift har behov for å overvåke deler av dataen, for å se etter spesielle tilstander, kan BI benyttes til dette. Dette gjør det mulig å respondere raskt ved tilstandsendringer. Det er mulig å skrive jobber som kjører kontinuerlig på BI-serveren, som overvåker et gitt datasett. Dersom et forhåndsbestemt mønster oppstår i dataen, eller en gitt tilstand inntreffer, reagerer jobben på dette. Reaksjonen kan være en varsling til gitte interessenter, i form av e-post, eller andre spesialtilpassede varslinger.

Dashboards gjør det svært enkelt for ledelsen å følge med på hvordan det går med bedriften på et overordnet nivå, ved å følge med på KPI-enes tilstander. Det kan for eksempel være et enkelt måle-termometer som viser hvordan bedriften ligger an i forhold til budsjett, som vist nedenfor.



Figur 2.10: Eksempel på dashboard som viser salg i forhold til budsjett.

Bedrifter kan også lage dashboards for mer “lav-nivå-informasjon”, for å hjelpe ansatte å utføre arbeidet sitt. Det tidligere eksempelet om bestillingsansvarlig i matbutikken kan brukes igjen her. Bestillingsansvarlig kan ha et dashboard på PC-skrivebordet som ser ut som en lampe. Lampen lyser grønt dersom ingenting behøves å bestilles, og blinker rødt dersom det snart er tomt for en vare på lager. Da blir bestillingsansvarlig oppmerksom på at han må sende inn bestilling på varen.

Mange bedrifter benytter seg av BI-dashboards for å overvåke ansattes prestasjoner, og gir bonuser basert på disse [71]. Dette bygger det psykologiske prinsippet om at man jobber hardere og mer effektivt dersom man vet at man blir overvåket. Dette, i kombinasjon med at man ønsker en høyere bonus, er ment å skulle gi de ansatte høyere motivasjon til å jobbe hardt. Et eksempel på slik bruk av BI, er en av Amerikas største linseleverandører, 1-800-CONTACTS [71], som har implementert en BI-løsning som bruker dashboards for å vise en oversikt over kundebehandleres prestasjoner. Denne oversikten er tilpasset hver enkelt kundebehandler, og viser statistikk over gjennomsnittelig samtalevarighet, gjennomsnittelig ordrestørrelse, prestasjoner den spesifikke dagen i forhold til en gjennomsnittelig dag den siste måneden, og hvordan deres prestasjonen ser ut sammenlignet med en gjennomsnittelig ansatt på avdelingen. Data som vises på dashboardene oppdateres hvert 15 minutt. Månedlig bonus er knyttet til disse metrikkene, og denne BI-løsningen har vist signifikant forbedring i kundebehandleres prestasjoner [71].

2.3.5 Strategi

En strategi defineres som en plan for å oppnå et gitt mål. Noen av de bedriftene som bruker BI, ønsker å benytte BI for strategi-føring. Eksempelet om det amerikanske linseselskapet, er et godt eksempel for BI-strategi; ved

å bruke dashboards oppnådde bedriften høyere effektivitet blant de ansatte. Et annet eksempel på vellykket bruk av BI for strategi, er gambling-firmaet Harrah's Entertainment i Las Vegas. Harrah's Entertainment transformerte hele sin konkurransestrategi ved å benytte seg av BI og data mining. De brukte data om kundene til å bygge et kundesentrert datavarehus, som lagret data om spillingen, hotellbruken og deltakelse på ulike arrangementer. Ved å analysere denne dataen fikk firmaet ny informasjon om kundenes preferanser, livstidsverdier, preferanser av ulike spill, og hvilke promoterte tilbud som var mest sannsynlig å bli akseptert i ulike markedssegment [71]. Harrah's Entertainment gikk fra å ha en intuisjons-basert forretningsstrategi, til å basere strategien sin på statistiske analyser og mønstergjenkjenninger. Denne BI løsningen har gitt stor ROI, og andre gambling-selskaper har kopiert deres forretningsstrategi.

2.3.6 Beslutningsstøtte

Ved å følge med på dashboards over KPI-er, og gjøre bedriftsanalyser, kan ledelsen i bedrifter få hjelp til å vite hva slags avgjørelser som er riktige å ta. BI som beslutningsstøtte er derfor blitt en viktig motivator for mange av de som implementerer BI. Beslutninger tas ofte på bakgrunn av analyser over ting som har hendt, men også over ting som ser ut til å kunne skje i framtiden. Ved å bruke data mining/mønstergjenkjenning kan bedriftene gjøre antakelser om hvordan mønsteret vil se ut i framtiden [67], og ta beslutninger basert på dette.

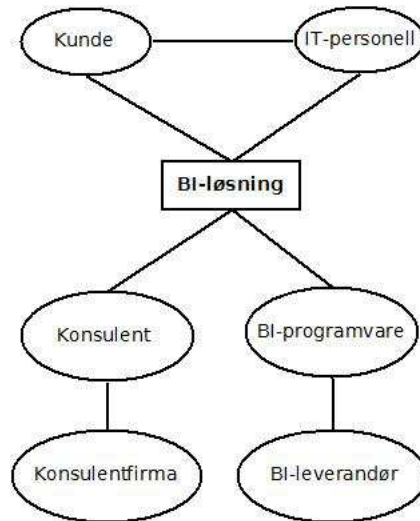
Ved å oppdatere datagrunnlaget hyppig, kan man gjøre beslutninger basert på relativt ferske data. Dette er ønskelig hos noen bedrifter, dersom beslutningen som skal tas er tids-kritisk. Et eksempel på bruk av BI som beslutningsstøtte, er løsningen til det Amerikanske flyselskap, Continental Airlines [71]. Løsningen kjører en beslutnings-analyse dersom et fly er forsinket, for å finne en optimal måte for håndtering av verdifulle passasjerer som skal videre med andre fly. Analysene baseres på data om kundeprøft, kundereservasjoner, real-time flydata fra flyet, ankomstdata, gate-beliggenhet, og avgangsdata til neste fly (tidspunkt og gate-beliggenhet). Basert på disse analysene regner de ut hvordan de på best mulig måte kan bringe verdifulle kunder og deres bagasje så effektivt som mulig til neste gate/fly innen flyavgang.

2.4 BI i praksis

Når en bedrift skal implementere en BI-løsning er det flere involverte parter. Det kreves både programvare for å få satt opp løsningen, og utviklere som kan gjøre selve implementasjonen.

Det finnes mange BI-leverandører på markedet, som leverer både proprietære- og Open Source BI-løsninger. En komplett BI-implementasjon krever programvare for design av database, ETL, rapportering, kube-bygging, analysing, aggregat-design, data-mining, visualisering osv. Mange leverandører leverer komplette plattformer som inneholder alle disse verktøyene, mens andre bare leverer enkeltstående verktøy. Det er tøff konkurranse blant BI-leverandørene, ettersom BI er et veletablert felt, med bred enighet om hva en god BI-plattform burde tilby [16]. I følge Eckerson kan dette gjøre det vanskelig for kunden å velge plattform.

I likhet med mange andre avanserte programvareinstallasjoner, krever BI-implementasjoner omfattende modifikasjoner og konfigurasjoner. Dette krever kompetanse som svært få bedrifter er i stand til å innhente internt. Det blir derfor ofte brukt eksterne konsulentfirma for å utvikle BI-implementasjonen. Konsulenter kan leies for å jobbe med hele, eller deler av løsningen, og det er noen ganger ønskelig å leie inn eksperter på spesifikke områder av implementasjonen (datavarehus, ETL, visualiserings osv). Noen av de største IT-konsulentselskapene, som Accenture og Capgemini [59], har egne kompetansesenter for BI-utvikling, mens andre selskaper, som Affecto, spesialisere seg på kun BI



Figur 2.11: Figuren viser involverte parter i en BI-implementasjon.

BI ble på grunn av de store lisenskostnadene for programvare lenge brukt av kun store bedrifter, men de senere årene har lisenskostnadene sunket, i tillegg til at det har kommet flere gode alternativer innenfor Open Source BI. Dette har ført til at flere mindre bedrifter også har hatt råd til å investere i BI. Dette vil diskuteres nærmere i Kap. 3.

2.4.1 utfordringer

En utfordring hos mange bedrifter, spesielt de mindre bedriftene, er mangel på intern IT-kompetanse. Selv om det er mulig å leie inn eksterne konsulenter som kan gjennomføre BI-implementasjonen, vil det være behov for interne ressurser som har tilstrekkelig teknisk kompetanse- gode nok til å lage rapporter og analyser.

Å hente ut reliable data fra ulike systemer, og kvalitetssikre disse, er alene en stor utfordring. ETL-steget blir ofte undervurdert av bedrifter som ønsker å implementere BI. I realiteten tar ofte ETL-steget opp mot 80 % av det totale implementasjonsbudsjettet [47], og står for 50 % av de uforutsette prosjektkostnadene [71].

BI-løsningen er avhengig av bedriftsbrukerens anerkjennelse, ved at de skjønner fordelene med BI, og ser mulighetene som eksisterer. Det er svært ofte motstandere blant kundens ansatte når en BI-løsning fører til endringer i bedriftsprosesser. Ansatte som tidligere hadde ansvar for data og prosesser føler at de mister kontrollen over noe som de har en sterk eierskapsfølelse

overfor. Dette gjør det til en utfordring å overbevise ansatte om at BI er en god idé. Derfor implementeres BI ofte iterativt, slik at ansatte gradvis selv ser at det er nyttig med BI. Denne tilnærmingen er mer eller mindre enn industristandard [47][16].

BI er bedriftsdrevet, og bedrifter som implementerer BI må derfor holde en bedriftsdrevet utviklingsmodell, i stede for en IT-drevet utviklingsmodell. Moss og Atre [47] hevder at IT-sentrerte utviklingsmodeller ofte går for sakte og dessuten er kompromissløse for bedriftsbrukeren. IT-utviklere, på sin side, ser ofte bedriftsbrukere som utålmodige og uforstående, med stadige endringer i kravene. Tett samarbeid mellom det bedriftsdrevne og det tekniske teamet er derfor svært viktig [16].

Utfordringer for mindre bedrifter er ofte knyttet til økonomi. For mindre bedrifter er ofte BI-programvare for dyrt, samtidig som det er vanskelig å regne ut ROI, da de ikke har et klart bilde over hva de vil få ut av BI-løsningen. Utfordringer for mindre bedrifter vil diskuteres nærmere i Kap 3.

Kapittel 3

OSBI

Open Source BI-programvare, heretter omtalt som OSBI, har det siste tiåret tatt opp kampen med de store, kommersielle BI-løsningene. Det vil i dette kapittelet bli sett på hvor langt OS-alternativene har kommet i utviklingen, hvor modne de er, og hvordan utviklingen kan komme til å se ut i framtiden.

Kapittelet starter med en seksjon om Open Source som utviklingsmetode, samt en sammenligning med andre typer programvare. Seksjon 2 tar for seg OSBI, og OSBI's modenhet, mens seksjon 3 til slutt tar for seg mulige framtidsutsikter for OSBI.

3.1 Open Source

Open Source¹, heretter omtalt som OS, er programvare med åpen kildekode, som betyr at kildekoden kan lastes ned gratis fra internett. Den vanligste lisensen er GPL²-lisensen [37], som sier at kildekoden skal være åpen, mulig å modifisere, og å gjendistribuere under samme lisens.

3.1.1 Open Source som utviklingsmetode

OS oppstod på 1960-tallet, som et resultat av at forskere som brukte datamaskiner til arbeidet sitt var avhengige av å dele kode seg imellom. Da kommersiell programvare senere ble tilgjengelig, var det ofte kun produsentene som hadde tilgang til kildekoden, og OS ble da et attraktivt alternativ, siden det tillot brukere å modifisere koden for å tilpasse ens personlige behov [30].

¹På norsk: åpen kildekode.

²General Public Licence

OS-prosjekter starter ofte av en enkelt person, eller en liten gruppe, som ønsker å utvikle noe som en selv skal bruke. Et godt OS-prosjekt er tilrettelagt slik at en utvikler fra hvor som helst i verden lett skal få til å sette seg inn i prosjektet, og raskt komme i gang med bidrag, i form av dokumentasjon og/eller kode [23]. Det blir dermed deres felles interesse å utvikle programvaren, noe som binder utviklerne sammen, og gjør de til et samfunn/fellesskap [77]. På mange OS-prosjekt møtes aldri utviklere ansikt-til-ansikt, men kommuniserer over internett. Grunnet denne typen kommunikasjon, er det ofte tilfelle at utviklere på samme prosjekt er fra ulike steder i verden.

Et OS-prosjekt kan bestå av alt fra noen få til flere tusen utviklere, og det finnes i dag over 260 000 OS-prosjekt [61]. Det er ikke uvanlig at et prosjekt legges dødt, fordi utviklerne som tidligere deltok ikke lenger har noe utbytte av å være med [77]. Den største risikoen ved investering i OS-teknologi er nettopp denne risikoen for at prosjektet kan legges dødt [28]. Det er også en fare for at prosjektet endrer lisens, noe som kan endre forutsetningene for framtidig bruk. Grunnet disse risikoen velger mange å ikke velge OS-programvare.

Det er ikke lenger slik at det kun er frivillige bidragsytere som utvikler Open Source-programvare; det har etter hvert blitt mer vanlig at bedrifter velger å bidra på disse prosjektene, ved for eksempel å sette egne ansatte på et allerede eksisterende OS-prosjekt, eller ved å gi direkte økonomisk støtte til prosjektet. Ved å gjøre dette, kan bedriftene få være med å avgjøre hvordan videre utvikling av programvaren skal gjøres.

3.1.2 Andre typer programvare

Motsetningen til OS-programvare er proprietær/lukket programvare. Leverandører av denne typen programvare ønsker ofte å selge den, og de må derfor begrense tilgangen til kildekoden, ved å bruke en lukket lisens. Programvare som selges kalles kommersiell programvare [32]. Det er viktig å se kilet mellom proprietær og kommersiell programvare, da kommersiell programvare også kan være Open Source.

- Proprietær programvare: Kildekode er lukket
- Open Source programvare: Kildekode er åpen
- Kommersiell programvare: Programvare selges
- Kommersiell Open Source: Programvare selges, og er basert på Open Source

Det finnes to typer kommersiell OS-programvare: 1) Programvare er basert på OS-programvare, men har proprietær programvare på toppen av denne, og 2) programvare har helt åpen kildekode, og gir leverandøren fortjeneste igjennom tilbud om support, dokumentasjon, og opplæring. Den mest utbredte lisensieringsmodellen som finnes i OSBI-markedet i dag er en kombinasjon av disse to, som beskrives i neste seksjon.

3.2 Open Source Business Intelligence

I årevis bestod BI-markedet utelukkende av kommersiell programvare [28]. Etter hvert begynte noen få enkeltstående OSBI-verktøy å utvikles. I starten bestod disse av isolerte deler for datavarehus-utvikling, med et begrenset sett med funksjonaliteter. Senere ble noen av disse enkeltstående verktøyene kjøpt opp/satt sammen til hele OSBI-plattformer, som da bestod av alle nødvendige moduler for å gjøre en komplett BI-implementasjon. Slike plattformer bestod av programvare for databaseutvikling, integrasjon, rapportering, og analysering. Den første plattformen på OSBI-markedet var Pentaho, som startet opp i 2004. Thomsen og Pedersen [66], og Golfarelli [28] mener at Pentaho er den ledende OSBI-plattformen på OSBI-markedet. Pentaho vil derfor være i fokus, når OSBI's modenhet senere skal vurderes. Andre store plattformer er Jaspersoft, SpagoBI, Actuate BIRT og Palo. Det finnes også en rekke mindre OS-prosjekter, som tilbyr enkeltstående OSBI-verktøy.

3.2.1 Lisensieringsmodell

Både Pentaho og Jaspersoft benytter en kommersiell OS-modell for lisensiering [28]. De tilbyr to versjoner: 1) en helt gratis OS-versjon, som dekker brukerens grunnleggende behov, og 2) en enterprise-versjon som kan kjøpes. Den sistnevnte fungerer som et medlemskap, som gir tilgang på support, dokumentasjon og opplæring. I tillegg har den siste versjonen noe ekstra funksjonalitet som holdes proprietær. Ved å holde en slik lisensieringsmodell har disse leverandørene klart å dra nytte av et utviklingsmiljø basert på frivillige utviklere, samtidig som de tjener penger på å gi betalende kunder flere fordeler. Når det i denne oppgaven snakkes om kommersiell BI, menes tradisjonell kommersiell BI, og ikke kommersiell OSBI. Den sistnevnte vil referes til som KOSBI³.

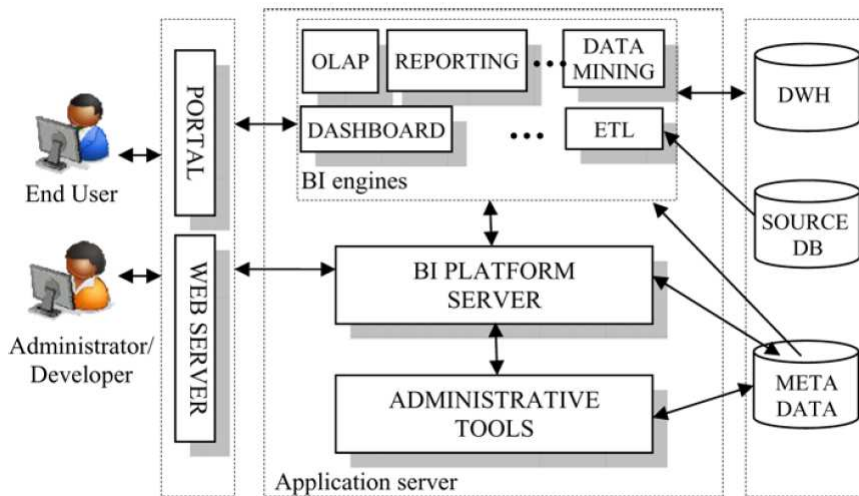
³Kommersiell Open Source Business Intelligence.

3.2.2 Arkitektur

OSBI-plattformene er satt sammen av ulike enkeltstående OSBI-verktøy, og får derfor en modulær arkitektur [28]. Pentaho har vokst fram ved at de har kjøpt opp/tatt over mindre prosjekter og lagt det under Pentaho. Jaspersoft har hovedsakelig benyttet seg av partnerskap med andre prosjekter, og på denne måten fått påvirket evolusjonen av programvaren.

De ulike OSBI-verktøyene/plattformene bruker ofte de samme standardene og filformatene. Kombinasjonen av åpne standarder og et modulært design, gjør at man kan kombinere verktøy fra ulike plattformer. Jaspersoft og Pentaho har dratt nytte av dette, ved å samarbeide om utvikling av noen av verktøyene som de bruker i sine plattformer.

OSBI-plattformene er utviklet i Java, ettersom modulene de bruker er basert på denne teknologien. De krever typisk en applikasjons-server som brukere og administratorer kan aksessere igjennom en web-browser. Arkitekturen vist i figuren nedenfor viser arkitekturen til Jaspersoft og Pentaho.



Figur 3.1: Arkitektur for Jaspersoft og Pentaho [28].

3.2.3 OSBI-markedet

En undersøkelse gjort på BI-markedet, i USA i 2010, viser at 48 % av OSBI-brukere har omsetning på mindre enn 50 millioner dollar, mens 37 % av brukerne av tradisjonelt lisensiert BI er like små [1]. Dette tyder på at omsetning i bedriften påvirker valget av programvare. En undersøkelse gjort

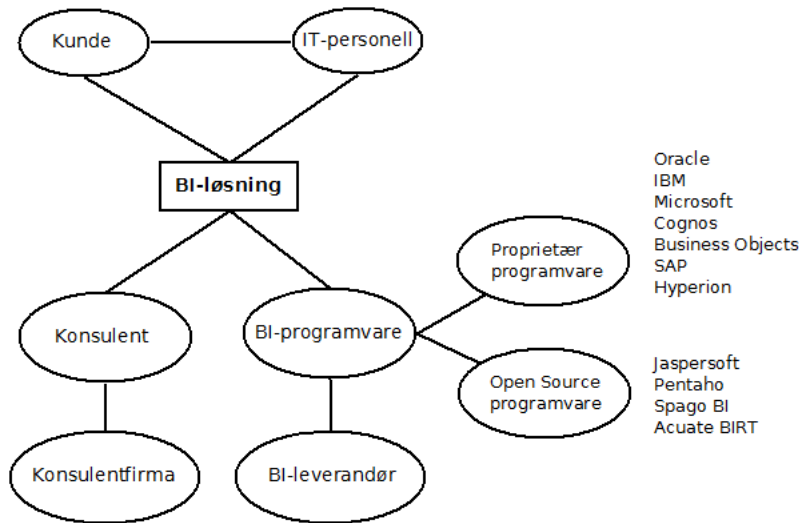
av Pentaho i 2008 viste at 80 % av de som benyttet seg av Pentaho ønsket å få mer ut av bruken, men 84 % mente også at for store kostnader var hemmende for disse planene.

I følge Golfarelli tiltrekker OSBI seg også forskere [28]. BI kan brukes for å prototype, teste, og analysere datagrunnlag som samles av forskere. Han mener også at mange offentlige instanser foretrekker OS-programvare, og at OSBI derfor kan bli en viktig del av IT for offentlig sektor. Gartner [7] er enig i at OSBI er interessant for offentlig sektor, og legger til at OSBI er spesielt attraktivt for mindre bedrifter.

Av geografiske lokasjoner har Øst-Europa og Asia utmerket seg, med stor interesse for OSBI [7]. En undersøkelse gjennomført av Gartner [7] i 2009, målte at 2 % av bedrifter som brukte BI, brukte OSBI.

3.2.4 OSBI Vs. tradisjonell BI

De store kommersielle BI-leverandørene tilbyr diverse integrerte pakke-løsninger, og skiller seg fra hverandre ved funksjonalitet, pris og brukervennlighet [32]. OSBI er en egen type produkt; det modulære designet gir følelsen av et eneste produkt, samtidig som de ulike delene kan brukes hver for seg. Den modulære arkitekturen og de åpne standardene, gir større fleksibilitet, noe som kan gjøre det lettere å gjennomføre integrasjon i heterogene IT-miljø.



Figur 3.2: Figuren viser involverte parter i en BI-implementasjon, og hvilke store BI-plattformer som eksisterer av proprietær BI og OSBI.

Fordeler ved bruk av OSBI:

- Billigere programvare
- Leverandøruavhengighet [32]
- Større fleksibilitet
 - Mulig å blande verktøy fra ulike plattformer
- Muligheter for å få delta i videre utvikling av programvare

Fordeler ved bruk av tradisjonell BI:

- Leverandør har vært på markedet lengre og tilbyr derfor en mer sofistikert programvare
- Et helhetlig produkt
- Tilgjengelig support
- Tilgjengelig dokumentasjon
- Ingen risiko for at BI-leverandøren skal legge prosjektet dødt

Gartner [7] mener at bedrifter som velger å bruke OSBI må ta dette valget basert på de samme beslutningskriteriene som de ville benyttet for valg av proprietær BI. Slike beslutningskriterier kan være funksjonalitet, reliabilitet, mulighet for intern support, og mulighet for tredje-part support (eksterne konsulenter). Ved å bruke de samme beslutningskriteriene, unngår man å velge OSBI kun fordi programvaren er billig/gratis. Gartner mener videre at bedrifter som ikke har signifikante ressurser til utvikling, og har risiko assosiert til programvareutvikling, burde la være å forsøke å forstå kildekoden, men heller kjøpe support der det er nødvendig.

3.2.5 Modenhhet

Gartner [7] spådde i 2008 at det ville være en tre-bølge-tilnærming til bruk av OSBI:

- 2004 til 2007: Tidlig bruk
- 2008 til 2012: Drevet av midt-markedet (SMB)
- fra 2012: Et vanlig valg

Kvaliteten på OS-programvare testes tradisjonelt av tre instanser som finansieres av EU [28]: Flossmetric⁴, Qualoss⁵, og SQOOS⁶. Disse har forsøkt å utvikle en standard metodologi for å teste kvalitet på OS-programvare. Ingen av de tre har testet kvaliteten på noen av OSBI-plattformene som finnes på markedet. Det er derfor sett på andre vurderinger gjort av OSBI. IT Novum [32], Thomsen og Pedersen [66], og Golfarelli [28] har hver for seg gjort vurderinger av noen av de største OSBI-tilbudene på markedet.

IT Novum [32] gjorde i 2009 en sammenligning av Jaspersoft, Pentaho og Palo. Det ble da konkludert med at alle tre plattformene tilfredsstilte typiske BI-scenarier, og at det var marginale forskjeller mellom plattformene. Pentaho utmerket seg likevel som den definitivt mest lovende av de tre. Det ble foreslått at mindre bedrifter som oftest vil ha nok med den helt åpne versjonen, mens større selskaper fint kunne ta i bruk kommersiell OSBI. Pengene spart på programvare kunne da heller brukes for å konfigurere BI-løsningen etter bedriftens spesifikke behov. IT Novum påpekte at den lave prisen ikke alene burde være avgjørende for valg av OSBI, da det er essensielt at valgt plattform passer bedriftens behov.

Golfarelli [28] gjorde i 2009 en funksjonell og arkitektur-sammenligning av Jaspersoft, Pentaho og SpagoBI. Golfarellig kartla da hva som kun ble tilbydt med kommersiell OSBI-versjon, og ikke med helt åpen OSBI-versjon:

- Bedre administrative konsoller
- Lettere konfigurering: Slipper manuell konfigurering
- Tillater prosessovervåkning: Gjør det lettere å optimalisere kjøretiden til prosessene
- ETL debugging GUI: Minsker tiden til utvikling betraktelig
- Større og bedre dokumentasjon og en kunnskapsbase

Golfarelli mente etter denne sammenligningen at helt åpen OSBI ikke kan brukes når man skal utvikle kritiske applikasjoner, som skal brukes i komplekse organisasjoner. Det vil da være mer fornuftig å bruke kommersiell OSBI, slik at man lettere kan administrere løsningen, og overvåke prosesser. Golfarelli konkluderer med at selv om OSBI ikke er like sofistikert som de store kommersielle plattformene, må de ansees som valide alternativer til

⁴Free/Libre Open Source Software Metrics.

⁵Quality in Open Source Software.

⁶Software Quality Observatory for Open Source Software.

tradisjonell BI. Dette er spesielt tilfelle for SMB-er, hvor datakvalitet og ytelse ofte ikke er like kritisk som hos de store, mener Golfarelli.

Thomsen og Pedersen gjorde i 2008 et studie hvor de undersøkte enkeltstående OSBI-verktøy, og vurderte modenheten til disse. Undersøkelsen var en videreutvikling av en undersøkelse de hadde gjort i 2004, altså 4 år tidligere⁷. I vurderingen hadde de fire grupper med verktøy: ETL, Databasehåndteringssystemer (DBMS), OLAP-servere, og OLAP-klienter. For hver gruppe så de på om følgende krav var oppfylt:

- ETL:
 - Må støtte et bredt spekter av kilde-systemer som kan integreres
 - Inkrementell lasting: det må være mulig å bare laste endringer i datagrunnlaget (ikke hele kubene på nytt)
 - Hvordan ETL-jobben lages: GUI eller XML-koding?
 - Mulighet for parallell kjøring av jobber (viktig når store datamengder skal lastes i datavarehuset, og tid er kritisk)
- DBMS:
 - Mulighet for håndtering av store datamengder
 - Mulighet for replikering og partisjonering
 - Programmeringsspråk som støttes
- OLAP-servere:
 - ROLAP eller MOLAP server?⁸
 - Størrelse på datakilder som kan håndteres
 - Underliggende DBMS som støttes
 - Spørre-språk mot server
 - Mulighet for å forhåndskalkulere aggregat-tabeller
 - Hvordan kube-generering gjøres: GUI eller XML-koding?
- OLAP-klienter:
 - Hvilke OLAP-servere denne kan brukes sammen med

⁷Undersøkelsen ble gjort samme år som den første BI-plattformen oppstod.

⁸Relasjonell eller multidimensjonal server.

- Hvilke spørrespråk som støttes
- Typer rapporter som støttes

Etter undersøkelsen gjort i 2004 konkluderte Thomsen og Pedersen med at det kun var databasehåndteringssystemet som kunne ansees som modent. Fra 2004 til 2008 var det skjedd store endringer, og de mente det hadde vært en sterk utvikling hos de store leverandørene som Pentaho og Jaspersoft. Det var inkludert flere verktøy i plattformene, og generelt mente de at verktøyene i plattformene var mer modne og kraftigere enn før.

3.2.5.1 Pentaho

ETL fikk gode vurderinger etter 2008-undersøkelsen. Det fantes nå et grafisk grensesnitt for design av jobber og transformasjoner. Inkrementell lastning var mulig, ved at Kettle (integrasjonsverktøyet i Pentaho) kunne logge jobber som kjørte, og dermed bruke tids-stempel i loggen for å kun laste ny data. Parallelle jobber var ikke støttet i den daværende Pentaho-versjonen, versjon 3.0.

Pentahos ROLAP-server, Mondrian, hadde fått bedre og raskere funksjonalitet. Denne var implementert i Java, og kunne derfor brukes sammen med de fleste DBMS, noe som økte fleksibiliteten. Serveren støttet de-facto standarden i BI-inustrien, MDX-spørringer mot kube. Kube ble laget ved hjelp av XML-koding, og hadde altså ingen grafisk grensesnitt for å gjøre dette. De anså derfor design av kube som lite brukervennlig. Support for forhåndskalkuleringer av aggregat-tabeller var tilgjengelig, noe som lettet utviklingsjobben betraktelig. Brukere rapporterte at ytelsen var god, selv med flere hundre gigabyte i omløp.

3.3 Framtidsutsikter

Gartner [7] mener at OSBI definitivt er her for å bli, og at det innen 2012 vil være et helt vanlig valg innen BI. Golfarelli [28] tror at OSBI vil kunne utvikles raskere enn kommersiell BI i framtiden. Dette begrunner han med at OSBI ikke er bundet av problemattikk rundt kompatibilitet, i tillegg til at de slipper å være bundet til en utdatert og lite fleksibel arkitektur. OS-løsninger har også en fordel med at det ofte er flere hundre frivillige utviklere som driver utviklingen framover. Golfarelli er likevel usikker i sine spekulasjoner, ettersom historien til disse produktene er såpass kort. Han mener det er umulig å si om plattformene vil tjene nok fra de tjeneste de tilbyr i enterprise-versjonene til å bli på markedet. I følge evolusjonære trender hos OSBI-

plattformene, vil de om noen år være like gode som tradisjonell BI, mener han. For å bli enda bedre enn tradisjonell BI, må de holde den lave prisen, få all den samme funksjonaliteten som de kommersielle, men også komme opp med nye innovative funksjonaliteter.

Kapittel 4

2: SMB og BI

BI har lenge vært en teknologi kun benyttet av de store bedriftene. Dette er i ferd med å endres, og de siste årene har SMB fått stadig større innpass i BI-markedet. SMB utgjør hoveddelen av norsk næringsliv, da mer enn 99 prosent av bedriftene i Norge klassifiseres som SMB [57]. Det vil i dette kapittelet bli sett på hvilke utfordringer SMB møter i BI-markedet.

4.1 SMB

SMB er en forkortelse brukt om Små- og mellomstore bedrifter¹. Hvor mange ansatte en bedrift kan ha for å klassifiseres som SMB, har det vært- og er det fortsatt- internasjonale uenigheter om [75]. Termen er i størst grad brukt i USA, hvor den brukes om bedrifter med mindre enn 500 ansatte. I Europa har det vært mindre tall knyttet til termen, og i 2005 bestemte EU [24, 57, 18] at bedrifter med mindre enn 10 ansatte skal klassifiseres som *mikro*-bedrifter, bedrifter med færre enn 50 som *små* bedrifter og færre enn 250 ansatte som *mellomstore* bedrifter. EUs definisjoner vil bli brukt i denne oppgaven, og SMB brukes dermed som en samlebetegnelse på bedrifter med mindre enn 250 ansatte.

4.1.1 SMB i Norge

Det var i 1995 registrert 159 799 bedrifter i Norge, hvorav 99,5% hadde mindre enn 100 ansatte [57]. Disse bedriftene står for i underkant av halvparten av antall årsverk i norske bedrifter. Store bedrifter, som bare utgjør 0,5 % av bedriftsmassen, sysselsetter nesten 1/3 av arbeidsstyrken.

¹Engelsk: SME (Small and medium enterprises).

4.1.2 SMB vs. store bedrifter

Sammenlignet med store bedrifter skiller SMB seg ut på flere området. De har lavere omsetning, og dermed mindre ressurser, både med tanke på antall ansatte og i forhold til driftsbudsjett [9]. Med få ansatte oppstår en flatere struktur², og det eksisterer i mindre grad klart definerte roller for de ansatte. Dette kan føre til at de ansatte har flere ansvarsområder, sammenlignet med ansatte i de store bedriftene, hvor de ansatte ofte spesialisere seg innenfor mindre områder. Med flatere struktur følger også mindre byråkrati, da det er færre instanser involvert i administrasjonen.

Grunnet lavere omsetning sammenlignet med de store bedriftene, har SMB mindre ressurser å tildele utprøving av nye teknologier. Når det vurderes å investere i nye teknologier, som for eksempel BI, blir korrekt utregning av ROI derfor mer kritisk for SMB, da det vil ha større påvirkning på bedriften dersom profitt/utbytte er lavere enn investeringen. SMB kan derfor sies å ta en større risiko enn de større bedriftene, når det gjelder å investere kapital i ny teknologi.

4.2 SMB og BI

Markedet er i konstant endring; leverandører streber etter å oppnå innovasjon med nye produkter, mens bedrifter ønsker å oppnå verdi [76]. I mange år var BI-markedet dominert av de store bedriftene, og BI-plattformene var tilpasset disse. For noen år siden oppstod en endring i BI-markedet; mange BI-leverandører forsøkte å ekspandere sine løsninger, for å bedre passe SMB-enes behov [3]. Dette gjorde de ved å lage alternative løsninger som var enklere (mindre), billigere, og raskere å distribuere [17]. I tillegg har de store Open Source BI-leverandørene ila. det siste tiåret skapt en helt ny lisensieringsmodell, som beskrevet i kap 2. Dette har resultert i at det har blitt enklere for SMB å benytte seg av BI, både fra et teknologisk- og fra et økonomisk perspektiv. Denne utviklingen har gjort at SMB i stadig større grad har fått innpass i BI-markedet.

SMB-er går inn i BI-markedet med ulike hensikter. All [3] mener at SMB har vist et enormt behov for integrasjon, da mange SMB-er produserer massive mengder informasjon, men tenderer til å mangle analytiske egenskaper, noe som hindrer de å gjøre gode forretningsavgjørelser. Wise [76] mener at rapportering og automatisering av manuelle prosesser er vanlige motivasjoner for SMB-er. Gartner [27] er enig med Wise, og peker på det faktum at

²Flat struktur: mindre sprik mellom vanlige ansatte og ledelsen.

SMB ofte sliter med ineffektive og utdaterte/umodne manuelle prosesser for informasjonshåndtering [27]. IT-ansvarlige bruker ofte tilnærmet all sin tid på å overføre data manuelt. Noe data kan ha kritisk leveringstid, og skal brukes videre av andre prosesser. Det vil da være ønskelig å vurdere å automatisere slike prosesser, for å effektivisere både tiden til de ansatte, og dataleveringstiden. Ved å benytte integrasjons- og rapporteringsfunksjonaliteten som BI-verktøy tilbyr, kan SMB lage slike automatiserte prosesser.

En undersøkelse gjort av BARC [5] i 2010 ga resultater som viste at respondenter fra mindre bedrifter var mer fornøyde med sine BI-implementasjoner, sammenlignet med respondenter fra store bedrifter. Det ble da spekulert i om dette kunne komme av at BI var såpass nytt for de mindre bedriftene, og at forventningene og kravene derfor var lavere. Wise [76] foreslår at SMB-er kanskje ikke vet hva mulighetene er, og at de derfor går inn i BI-markedet med små ambisjoner, i forhold til hva slags muligheter som finnes ved bruk av BI.

4.3 utfordringer

Kernochan [34] skriver i sin artikkel fra 2010 at mye av BI-programvaren som eksisterer på markedet i dag, er spesialtilpasset SMB. Han mener likevel at denne programvaren ikke er moden nok til å kunne sikre SMB-er ROI, men at dette vil endres ilar. noen år, med enda bedre tilpasset programvare. Sett bort i fra umoden og dyr programvare, ser SMB-er på mange andre potensielle utfordringer relatert til bruk av BI [76].

Gartner [27] peker på fire hovedhindringer som SMB-er må overkomme for å lykkes med BI:

- Ansatte
- Prosesser
- Politikk
- Prioriteter

4.3.1 Ansatte

Wise [76] og Gartner [27] er enige om at IT-ansvarlige i SMB-er ofte er overlastet med arbeid, og derfor ikke har kapasitet til å få en BI-løsning som skal driftes i tillegg. Wise peker også på det faktum at de ofte ikke har god nok kompetanse om BI til å gjøre en komplett BI-implementasjon. Bedriftene har

teknologien og interaktiviteten tilgjengelig, men opplever et pedagogisk sprik mellom teknologien og de ansatte [76]. Wise tror at leverandører fokuserer for mye på teknologisk utvikling, og for lite på opplæring av utviklere. Det kan derfor bli vanskelig for utviklere som ikke har vært borti BI før å sette seg inn i det, kun basert på programvare-dokumentasjon. Gartner [27] mener at man i slike tilfeller er nødt til å leie inn ekstern BI-kompetanse.

Mange SMB-er er ikke villige til å leie inn eksterne konsulenter/rådgivere, og bruker kun egne ansatte. Det som ofte da er tilfellet, er at de mangler tilstrekkelig kompetanse om BI [20]. Dette kan gå galt på flere plan, da kompetansen må være til stede både på forretningsnivå og på teknisk nivå; Man må være i stand til å regne ut et godt estimat på ROI for BI-implementasjon, man må ha kompetanse om hvordan utviklingsmodell som burde benyttes, samtidig som man trenger teknisk kompetanse for å gjøre selve implementasjonen.

I følge Fisher [20] har SMB mindre rom for feil, sammenlignet med de store selskapene. Det er mer kritisk at utregning av ROI for en BI-implementasjon er korrekt, da SMB har mindre ressurser å falle tilbake på. For å kunne si noe om ROI av en BI-implementasjon, må man blant annet vite følgende:

- Hvor mye kapital den nødvendige programvaren vil koste
- Estimat over implementasjonstiden for BI-løsningen
- Hva man får igjen for implementasjonen i ren bedriftsverdi
- Estimat på hvor mye support som vil bli nødvendig på sikt
- Framtidige muligheter for utvidelse

Mange SMB-er besitter ikke god nok kompetanse til å lage slike estimater, og ender derfor opp med å feilberegner disse. En vanlig “felle” å falle i er at man har for stort fokus på målet, og ikke den faktiske veien fram til dette målet, mener Wise [76]. Grunnet mangelfull kunnskap om hva en BI-implementasjon innebærer, blir gapet for stort mellom hva som ønskes oppnådd og det som faktisk må gjøres for å oppnå dette. Man vil i slike tilfeller få et dårlig utgangspunkt for implementasjonen, da man får et urealistisk bilde av veien fram mot en ferdig BI-løsning. Dette kan resultere i en BI-implementasjon som blir langt dyrere enn antatt. Wise [76] eksemplifiserer dashboards som en funksjonalitet som lokker mange SMB-er til å investere i BI. De ønsker å bruke dashboards for å måle gitte bedriftsmetrikker, og få varslinger når disse endres, men de er ikke klar over alle lagene som må ligge i grunnen for å få laget et slikt dashboard.

Noen SMB-er velger å bruke ressurser på å utdanne egne ansatte som skal jobbe med bedriftens BI-løsning, for å være bedre rustet til å gjøre en god implementasjon. I slike situasjoner vil de ansatte typisk vite “mye om alt” rundt implementasjonen. Til sammenligning har ofte de store bedriftene ansatte som har spesialkompetanse³ på ulike områder innenfor BI. De vil da besitte mye kunnskap om et lite område, i stede for litt kunnskap om alt, slik som ansatte i SMB ofte gjør[76]. Wise mener at de store bedriftene på grunn av en slik “industri-utdanning” er kommet svært langt innenfor BI, sammenlignet med SMB. En problematikk rundt den typen BI-kunnskap som SMB opparbeider seg, er at de ansatte blir svært verdifulle. De vil besitte dyp kunnskap om både BI som teknologi, og bedriftens egen BI-løsning. Kompetanse- og kunnskapsoverføring vil i slike tilfeller være svært kritisk.

Gartner [27] foreslår som mulig løsning at man forsøker å sette sammen et utviklingsteam bestående av en kombinasjon av interne og eksterne ressurser. Ekspertise i analyse og intern arkitektur burde ideelt komme fra interne ressurser, mens teknisk kompetanse kan leies fra eksterne konsultentselskaper. Chen [10] mener at SMB-er får både en kortsiktig og langsiktig gevinst ved å leie konsulenter, da disse hjelper ansatte å forstå muligheter innen BI på lang sikt.

Et annet problem relatert til de ansatte, er at de ulike enhetene/avdelingene hos noen bedrifter kniver om budsjett, med ulike ønsker og prioriteringer. Dette, i kombinasjon med at noen ansatte er uvillige til å dele informasjon fra systemer de har ansvar for, bidrar til å skape en dårlig bedriftskultur for samarbeid. Gartner [27] mener at bedriftskulturen i slike tilfeller må endres fra å la ulike enheter/avdelinger holde data i siloer, og heller oppfordre disse til å samarbeide mot et felles datagrunnlag og et felles mål. En mulig løsning er å la minst én representant med ansvar for hver datakilde delta i konstruksjonen av infrastrukturen av BI-løsningen. Ved å samarbeide på en slik måte vil man unngå at enheter i samme bedrift kniver om budsjettet, og heller forsøker å oppnå langsiktig bedriftsverdi, ved å investere i en felles løsning for datalagring. Gartner foreslår også at det kan holdes en proaktiv kommunikasjon med de ansatte involvert, for å justere forventningene deres, og forberede de på at innovative endringer som vil komme. Dette gjelder både ansatte som vil miste arbeidsoppgaver (som følge av automatisering), og ansatte som i større grad må dele data fra sine systemer. På denne måten får de ansatte mulighet til å gradvis forberedes på et samarbeid, slik at de kanskje er mer positivt innstilte når tiden for endringene kommer.

³For eksempel en ETL-ekspert eller en datamodellerings-ekspert

4.3.2 Prosesser

Gartner [27] mener at automatisering av manuelle prosesser er en av de største motivasjonsfaktorene for implementasjon av BI hos SMB. Men dette er ressurskrevende, og ROI burde derfor vurderes for hver enkelt prosess, for å se på hvilke prosesser det vil lønne seg å automatisere [27]. Som nevnt tidligere, er utregning av ROI en svært omfattende prosess, og mange SMB-er er ute av stand til å beregne denne, selv om de får ekstern hjelp.

Å beregne ROI av en prosess er ingen enkel utregning, da det er mye som spiller inn:

- Prosessens viktighet for bedriften som helhet
- Ressurser involvert (innsats som kreves for dataforberedelser)
- Hvordan dataen/informasjonen vil konsumeres videre av andre prosesser og analyser [27].

Gartner [27] mener at bedriften burde sette opp en prioritetsliste, hvor hver prosess vurderes, og får en prioritet. Det er da viktig å ta i betraktning hvordan de prosessene som er relatert til den samme informasjonen, samhandler med hverandre, og hvordan data må struktureres for å møte leverings- og konsumeringskrav.

4.3.3 Politikk og Prioritet

Gartner [27] mener at SMB har en stor fordel med at ledelsen ikke er like stor som hos de store bedriftene. Færre i ledelsen betyr færre involverte parter i byråkratiske avgjørelser. Dette gir muligheten til rask respons på tilstandsendringer i bedriften, og bedriften vil typisk være smidig⁴. Men denne smidigheten er ikke gratis; bedriften endres kanskje så raskt, og så ofte, at det setter opprettholdelsen av konsistente og reliable bedriftsmetrikker på spill. Det kan bli vanskelig å lage standardiserte, konsistente målinger til datagrunnlaget, når dataen hele tiden endres. Uten konsistent, pålitelig bedriftsinformasjon, og gode analytiske prosesser, er det umulig for bedrifter å gjøre bedriftsavgjørelser basert på BI. Informasjon spiller en strategisk rolle for å tilrettelegge for at en forretning skal kunne respondere rasjonelt og effektivt på endringer i bedriften.

Gartner [27] og Kernochan [34] er enige om at et viktig kriterium for at en SMB-er skal bruke BI optimalt, er at de utnytter fleksibiliteten/smidigheten

⁴En smidig bedrift er en bedrift som enkelt kan endre forretningsstrategi

til det fulle. For å få til dette mener Gartner [27] at SMB-er trenger fleksibilitet i informasjonshåndteringen, og følgende punkter foreslås for å sikre en slik fleksibilitet:

- Det må dannes klare forståelser om forventningene til levering av informasjon
- Det må lages prioritetsnivå for dataen som skal håndteres
- Det må lages en plan for bedriftsstrategien framover
- Bedriften må være innstilt på å lære av sine feil
- Det må holdes en kontinuerlig kommunikasjon mellom alle involverte parter
- Ledelsen må være åpen for endringer i planen, samt skiftende bedriftsprioriteter

Fleksibiliteten gjelder ikke bare på kort sikt, men også på lang sikt; det må tilrettelegges for inkrementell forretningsbygging. Gartner [27] foreslår at SMB-ene burde bruke en tilnærming til BI hvor de tenker stort, men starter smått, for å holde fokus på en solid grunnmur. Dermed tilrettelegges det for framtidig utbygging av implementasjonen. Kernochan [34] mener at nesten alle SMB-er har en “ønskeliste” over ting de ønsker å oppnå på sikt. De burde derfor lage en liste over disse, med prioriter. Dermed vil disse være klare, når bedriften i framtiden har samlet tilstrekkelige mengder historiske data, og dermed kan gjøre ønskelige analyser.

4.3.4 Oppsummering

Gartner [27], Wise [76], Kernochan [34], Fisher [20] og All [3] har flere ulike syn på hva som er utfordringer for SMB-er ved implementering av BI. Oppsummering av disse:

- Høy Risiko
 - Begrensede økonomiske ressurser
 - * Programvare for dyrt
 - ROI er kritisk
- Mangel på kunnskap om BI

- Mangelfullt bilde over muligheter
- Feilestimering av ROI
- Mangel på teknisk kompetanse
 - Urealistisk bilde over implementasjonsprosessen
 - Mangel på interne utviklere med kompetanse for utvikling og vedlikehold/drifting
- Motvilje fra ansatte
 - Føler eierskap til data, og ønsker ikke å dele denne
 - Mangel på samarbeid
 - Ulike prioriteter
 - Kniving om budsjett

Kapittel 5

Empirisk metode

Empirisk forskning er en vitenskapelig forskningsstrategi, som brukes for å besvare empiriske forskningsspørsmål. For å besvare disse spørsmålene benyttes *empirisk metode*. En empirisk metode brukes for å samle og analysere kvalitativ eller kvantitativ data [74]; det er en metode for å *oppnå kunnskap*. Hvordan kunnskap man ønsker å oppnå, avgjør hvordan forskningsmetode man burde benytte.

5.0.5 Kvantitativ vs. Kvalitativ metode

Kvantitativ og kvalitativ metode er de to hovedkategoriene for forskningsmetoder, og begge disse kan benyttes for å samle og analysere data. Det finnes ikke et rett eller galt svar på hva som er den beste av disse to, da dette er avhengig av konteksten metoden benyttes i. Bickman og Rog [6] mener at man først må se på spørsmålene man ønsker å besvare, og basere valg av metode på disse. Kvalitativ forskning tenderer til å besvare spørsmål på formen “*hva?*”, “*hvorfor?*” og “*hvordan?*”, og baserer analysen på *ord og drøftninger*. Denne metoden kan benyttes for å få en dybdeforståelse for et tema, og undersøker holdninger, oppførsel og erfaringer. Kvantitativ forskning besvarer spørsmål på formen “*hvor mange?*” og “*hvor ofte?*”, som samler data i form av nøyaktige *tall og statistikker* til analysen [60]. Denne metoden gjør det mulig å se på forhold mellom direkte målbare variabler.

Begge metodene har styrker og svakheter, og det er ofte ønskelig å bruke metodene komplementære til hverandre. I følge Clarke og Dawson [11] vil en forskers resultater være mer pålitelige dersom det er tatt i bruk ulike forskningsmetoder, da metodene vil styrke hverandre, ved å skygge over hverandres svakheter.

Kvalitative metoder:

- Etnografisk metode/observasjon
- Case-studie
- Semistrukturert og ustrukturert intervju
- Fokusgruppe

Kvantitative metoder:

- Spørreundersøkelse
- Strukturert intervju

Ved observasjon som metode, samles data gjennom å observere en situasjon, eller ved direkte deltakelse i situasjonen [49]. Deltakende observasjon krever at forsker er en lengre periode “i feltet”, for å skulle kunne legge fram detaljert dokumentasjon fra observasjonen. Observasjon som metode har blitt stadig vanligere å benytte for forskning på IT-systemer de siste 10-15 årene [13], da forskere ser verdi i den detaljerte dataen som er mulig å samle. Ved bruk av observasjon som forskningsmetode får forskeren et detaljert innblikk i hvordan organisasjonen som observeres fungerer, og hvordan de ansatte jobber sammen [63]. Forskeren får da et intimt kjennskap til dilemma, frustrasjoner, rutiner, forhold og risikoer som er en del av hverdagslivet til de ansatte. En ulempe med denne metoden er mengden tid som kreves for selve gjennomføringen, og for analyseringen av innsamlet data.

Case-studier samler, i likhet med observasjon, data om en situasjon, men da i ettertid av at denne fant sted. I et case-studie er primærdatakilde intervjuer, med dokumentasjon (f.eks. rapporter og møtereferat) som supplement. Et case-studie er mindre tidskrevende enn observasjon, men gir ikke like dyp forståelse.

Et intervju kan gi både kvalitative og kvantitative data, avhengig av hvordan intervjuet struktureres. Ved semi-strukturert og ustrukturert intervju bruker forsker et ukomplett skript med spørsmål, og det er nødvendig med improvisasjon fra forsker. Denne typen intervju er best egnet når man har et tema man ønsker å diskutere, for å gå i dybden på holdninger, tanker og meninger hos intervjuobjektet. Slike åpne intervjuer gir kvalitative data. Ved strukturert intervju er det ikke nødvendigvis forskeren selv som gjennomfører intervjuet, da det kun blir rom for å svare på gitte alternativer. Denne metoden brukes oftest i form av spørreundersøkelser på telefon, og dataene som samles blir da kvantitative. Intervju-metoden har noen vanlige fallgruver som forsker må være oppmerksom på. Disse er i hovedsak

relatert til konteksten og omgivelsene til intervjuet; Den kunstige situasjonen, tidspress og prestasjonsangst er noen av faktorene som kan påvirke intervjuobjektets svar [50]. En annen ulempe ved intervju er at det ofte er tidskrevende, da både forberedelser og etterarbeid kan ta mye tid.

Spørreundersøkelse som forskningsmetode gjør det mulig å samle inn data fra mange respondenter raskt, enkelt og billig. Svakheter med denne metoden er at den ikke går i dybden på spørsmålene, men som regel kun har avkrysning for rett svaralternativ. Det er også vanskelig å forsikre seg om at svarobjektene har skjønt spørsmålene korrekt.

5.1 Valgt tilnærming

Forskningsspørsmålene som diskuteres i denne oppgaven dreier seg om OS-BI's modenhet, og bruk av BI hos SMB. For å kunne forske på problemstillinger relatert til bruk av BI hos SMB, må man få en dybdeforståelse på faktisk bruk av BI hos SMB-er i næringslivet. Det er her spørsmål av typen "hva?", "hvorfor?" og "hvordan?", og passende forskningsmetode ble derfor kvalitativ metode.

Det ble spesielt tilrettelagt for å gjennomføre et feltstudie, med deltakende observasjon av en BI-implementasjon hos en SMB. Dette feltstudiet ble gjennomført hos konsulentselskapet Acando. BI-implementasjonen var i utgangspunktet en sommerjobb, og ble etter avtale utvidet til å utføres som et feltstudie. Feltstudiet ga detaljert innsikt i utviklingsprosessen; rutiner, utfordringer, dilemma, frustrasjoner og konflikter under veis.

Etter endt feltstudie hadde jeg oppnådd flere hypoteser som var ønskelig å teste, ved å sammenligne funnene fra implementasjonen hos Acando, med BI-implementasjoner hos andre SMB-er. Det var også ønskelig å se på forskjeller mellom SMB-ers implementasjoner og store bedrifters implementasjoner, for å se hvordan SMB-er skiller seg fra store bedrifter ved bruk av BI. For å kunne samle og sammenligne data fra mange, var det passende med en kvantitativt metode. Etersom spørreundersøkelser er lett å distribuere til mange, falt valg av metode på spørreundersøkelse. Denne ble distribuert til flere norske konsulentfirma som har erfaring med BI-utvikling.

Under analysen av spørreundersøkelsen oppstod flere resultater som var ønskelig å undersøke videre. Det ble da et behov for å gå i dybden av områdene for disse resultatene, noe som bare kunne gjøres ved å komme i dialog med noen som hadde erfaring på disse områdene. Intervju ble et naturlig valg, og det ble derfor gjennomført intervjuer av konsulenter som jobber med BI. Disse intervjuene resulterte i kvalitative data, som kunne

brukes som supplement til resultatene fra spørreundersøkelsen.

Hver metode ble utgangspunkt for neste metode brukt; hypotesene som ble generert under feltstudiet ble testet i spørreundersøkelsen; deretter ble resultatene som fra spørreundersøkelsen brukt som utgangspunkt for spørsmål til intervjuene. Ved å benytte en kombinasjon av kvantitativ og kvalitativ metode, er det gjort et forsøk på å skygge over svakhetene til de ulike metodene, og heller la de styrke hverandre, slik som Clarke og Dawson [11] foreslår.

Feltstudie → Spørreundersøkelse → Intervju

Forskningmetode	Tid brukt
Felt-studie	6 uker
Spørreundersøkelse	3 uker
Intervju	4 timer

Figur 5.1: Figuren viser tid brukt på de tre hovedmetodene brukt i dette studiet.

5.2 Datainnsamling

5.2.1 Tilgang

Ved datainnsamling er man avhengig av å ha tilgang til datakilder. Som sommeransatt hos Acando, fikk jeg et svært godt utgangspunkt for å gjøre et feltstudie på en BI-implementasjon hos en SMB. Jeg var da en av to utviklere, og fikk god innsikt i alle aspektene av implementasjonen. Da det senere ble holdt en spørreundersøkelse, var det nødvendig å finne respondenter innenfor målgruppen, som var konsulenter som jobber med BI. Noen av de som responderte på spørreundersøkelsen, ble funnet ved å gå på rekrutteringsdagene “Karrieredagen” og “itDAGENE” ved NTNU, hvor flere konsulent-bedrifter med BI-kompetanse var tilgjengelige. De resterende respondentene ble funnet ved å søke etter målgruppen på internett, og ved å få kontaktinformasjon fra bekjente. Mens spørreundersøkelsen pågikk, var det flere som annonserte interesse for den ferdige masteroppgaven. Noen av disse ble kontaktet i ettertid, med forespørsel om å delta på intervju.

5.2.2 Feltstudie



Figur 5.2: Figuren viser feltstudie, som den første av de tre metodene brukt.

Feltstudiet på Acandos BI-implementasjon, ble gjennomført som en kombinasjon av et case-studie og deltakende observasjon. Det ble også holdt en fokusgruppe med administrasjonen hos bedriften, samt intervju av oppdrags giver for prosjektet.

Feltstudie	Tid brukt
Case-studie	1 uke
Deltakende observasjon	6 uker
Fokusgruppe	2 timer
Intervju	1 time

Figur 5.3: Figuren viser hvilke forskningsmetoder som ble brukt ila. feltstudiet.

Noe av implementasjonen var gjennomført før jeg begynte å observere. Det ble samlet data om utviklingen, ved å bruke case-studie-metodikken. Det ble sett på rapporter og annen dokumentasjon laget av utviklerne, for å få innsikt i hva som var blitt gjort.

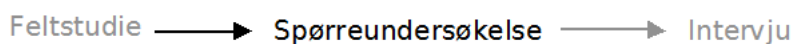
For hver dag ble det ført dagbok etter endt “jobbtid” for sommerjobben. I disse notatene ble det skrevet ned observasjoner, inntrykk, følelser, og anelser. Myers [49] påpeker at man etter å ha jobbet på et prosjekt en stund vil ta noen ting for gitt, og at det derfor er viktig å notere helt fra starten av. Myers anbefaler etnografiske forskere å sammenligne hva man tenkte i starten, med hva man tenkte på slutten av utviklingen. Dette skal bidra til å se hvorfor og hvordan det som var merkelig og interessant i starten, endret seg etter hvert. For å få til dette er det viktig med jevnlig refleksjon over det som er observert. Jeg forsøkte å følge Myers anbefalinger, og oppsummerte og drøftet dagboknotatene etter hver arbeidsuke.

Det ble ved én anledning gjennomført en fokusgruppe med administrasjonen hos bedriften, for å snakke om hva de var fornøyde og missfornøyde med etter første leveranse, og for å finne ut hva de så for seg å få ut av BI-implementasjonen. Det kom da flere gode innspill til forbedringer som kunne gjøres. Fokusgruppemetodikken var nyttig for å la deltakerne i gruppa diskutere, og gi hverandre idéer til mulige forbedringer.

Etter endt sommerjobb ble det holdt et intervju av oppdragsgiver, for å finne ut hvordan han mente implementasjonsprosessen hadde gått. Ved dette intervjuet ble Myers' og Newman's [50] dramaturgiske modell brukt, for å legge til rette for rett setting for intervjuet, og for å unngå vanlige fallgruver for bruk av intervju-metodikken. Intervjuet ble holdt semi-strukturert, for at intervjuobjektet skulle føle seg fri til å snakke rundt spørsmålene, og dermed få muligheten til å gå i dybden av tema som ble diskutert. Han ble informert om at jeg hadde "forsker-hatten" på, og at han måtte glemme min rolle som tidligere sommeransatt i bedriften. Det ble presisert at intervjuobjektet måtte forsøke å være ærlig, og snakke om det som var negativt, så vel som det som var positivt ved utviklingsprosessen. Han hadde fått spørsmålene på e-post dagen før, slik at han skulle få mulighet til å forberede seg. Intervjuet ble tatt opp ved hjelp av opptaksfunksjon på mobiltelefon, og etter intervjuet var ferdig, ble opptaket transkribert.

5.2.3 Spørreundersøkelse

I løpet av feltstudie fikk jeg et godt innblikk i hva en BI-implementasjon typisk innebærer. Jeg ønsket å få et mer generelt bilde over BI-markedet, og gjennomførte derfor en spørreundersøkelse, for å samle inn kvantitative data. Spørreundersøkelsen ble laget på engelsk, og distribuert ved hjelp av en http-link [62], for at flest mulig skulle ha mulighet til å svare.



Figur 5.4: Figuren viser spørreundersøkelsen, som nummer to av de tre metodene brukt.

5.2.3.1 Utforming av spørsmål

I utformingen av spørreundersøkelsen ble flere faktorer tatt med i betraktning. Et hovedmål var at den skulle være enkel å besvare, men likevel hente inn så mye verdifull data som mulig på 5-10 minutters svartid. Spørreundersøkelsen ble derfor gjennomført med lukkede spørsmål, med mulighet for å fylle inn kommentarer dersom man ønsket det. Målet med dette var at de som ønsket det kunne bruke lengre tid på å gi fylldige svar, men at de som ønsket å gjennomføre den raskt, skulle få muligheten til det.

Oates [52] regler for utforming av spørsmål i en spørreundersøkelse ble fulgt:

- Spørsmål burde inneholde maks 20 ord

- Alle spørsmålene må være relevante for den helhetlige spørreundersøkelsen
- Spørsmålene må være konkrete og entydige
- Ordvalgene i spørsmålene må være objektive

For at relevansen for hver enkelt respondent skulle være optimal, ble det laget flere stier i spørreundersøkelsen.

5.2.3.2 Pilottesting

Det ble gjennomført pilottest på to testobjekter, for å måle tiden det tok, og for å få tilbakemeldinger på kvaliteten på spørsmålene, og det helhetlige inntrykket av undersøkelsen. De to pilottesterne hadde ulik bakgrunn, og begge hovedstiene i undersøkelsen ble pilottestet.

Testobjektene ble instruert å gjennomføre undersøkelsen så raskt som mulig, for å få en minimal ca-tid på undersøkelsen, slik at så mange som mulig ville ta seg tid til å svare på undersøkelsen. Resultatene fra pilottestene ble benyttet i siste iterasjon av utvikling av spørreundersøkelsen, og antatt svartid ble regnet ut ifra hvor lang tid pilottesterne brukte. Testobjektene brukte i snitt 7,5 min., og estimert svartid ble derfor satt til 5-10 min.

5.2.3.3 Distribuering av undersøkelsen

Å finne mulige respondenter ble en utfordring, da BI fortsatt er relativt lite brukt i Norge. Kontaktinformasjon som ble brukt, var en kombinasjon av visitkort¹, informasjon funnet på bedriftenes hjemmesider, samt kontaktinformasjon mottatt av bekjente.

Potensielle respondenter fikk forespørsel på e-post om å delta på undersøkelsen. I denne forespørselen ble det informert om spørreundersøkelsens formål, at undersøkelsen var anonym, og at de bedriftene som ønsket det ville få en takk i forordet av oppgaven. Målet med løftet om anonymitet, var at de som ble forespurt ikke skulle være redde for å oppgi informasjon om sine kunders BI-implementasjoner og forretningsavgjørelser. Det ble hovedsakelig sendt e-post til ansatte med høyere posisjoner, med forespørsel om videredistribuering.

De som svarte at de var interessert i å delta, fikk tilsendt en http-link på e-post. Denne linken førte til spørreundersøkelsen. Fordelen med å la

¹Visitkort var samlet inn på stands på Karrieredagen og IT-dagene ved NTNU.

undersøkelsen ligge på en http-side, er at man unngår geografiske begrensninger, og dermed lett kan distribuere til respondenter over hele verden [52].

Andre fordeler er at det er billig og enkelt. Ulempen med bruk av http-link er at man ikke har kontroll på hvem som faktisk svarer.

Første side av undersøkelsen er vist i skjermbildet nedenfor. Her presenteres formål av undersøkelsen og antatt tid for gjennomførelse. Det ble forsøkt å gjøre utseende så enkelt som mulig, for at de som mottok linken skulle ta undersøkelsen seriøst, og dermed være mer sannsynlig å svare.

Introduction



This survey is conducted by a Master student at NTNU (Norwegian University of Science and Technology).

The purpose of the survey is to collect empirical data regarding the need for- and development of Business Intelligence solutions, and is therefore most relevant for technical BI developers.

The survey will take 5-10 minutes to complete, and your answers will be kept anonymous.

Thank you!

Next >>

Figur 5.5: Introduksjonen til spørreundersøkelsen.

5.2.4 Intervju

Resultatene fra spørreundersøkelsen ga indikasjoner på noen mulige trender i BI-markedet. Det var ønskelig å finne ut hvorfor disse trendene var som de var, og samtidig få et innblikk i ulike holdninger til OSBI. Det ble derfor gjennomført semi-strukturerte intervjuer, for å samle empirisk kvalitative data om tema som var av interesse.



Figur 5.6: Figuren viser intervju, som den siste av de tre metodene brukt.

10 respondenter av spørreundersøkelsen hadde meldt interesse for å få lese den ferdige masteroppgaven. 3 av disse fikk forespørsel om å delta på intervju, hvorav samtlige var villige til å delta. 3 åpne spørsmål ble sendt til intervjuobjektene på e-post, dagen før intervjuene. Intervjuene ble gjennomført med web-kamera og mikrofon, ved hjelp av verktøyet Skype. Notater ble ført under veis, og disse ble skrevet i detalj etter at hvert intervju var fullført. Intervjuene tok i gjennomsnitt ca 45 minutter.

Myers og Newman [50] påpeker noen vanlige fallgruver ved bruk av intervju-metodikken. Disse ble tatt hensyn til i gjennomføringen av intervjuene:

- *Mangel på tillitt:* Når forsker er en fremmed for intervjuobjektet, kan intervjuobjektet være skeptisk til å dele alle tanker med forskeren som holder intervjuet. I dette tilfellet kjente jeg ikke noen av intervjuobjektene, men de visste at jeg var student. Forhåpentligvis bidro dette til et mer avslappet forhold til situasjonen.
- *Mangel på tid:* Tidspress under intervjuet kan føre til at intervjuobjektet finner opp meninger under veis (uten at disse meningene var spesielt sterke, eller tenkt nøye igjennom). Denne fallgruven ble forsøkt håndtert ved at intervjuobjektene fikk spørsmålene på e-post dagen før, slik at de på forhånd kunne drøfte temaene, og gjøre opp seg opp reflekterte meninger rundt disse. Det ble også informert om at det ikke var noen tidsramme for intervjuet, for å unngå at intervjuobjektene skulle oppleve et tidspress.
- *Status-partisk:* En forsker kan noen ganger gjøre den feilen at han bare intervjuer personer med en viss status, og derfor misslykkes i å få det helhetsbilde som søkes fra flere parter. Personer med ulik status vil ofte ha ulike meninger og holdninger, noe som er viktig å ta i betraktning når man holder åpne intervjuer [50]. I dette tilfellet var det fire statuser av interesse; de som har erfaring med OSBI, de som kun har erfaring med kommersiell BI, de som har implementert BI for SMB, og de som har implementert BI for store bedrifter. Det var ønskelig å finne intervjuobjekter som dekket alle disse statusene. 2 av de 3 intervjuobjektene implementerte hovedsakelig OSBI-løsninger for SMB, men hadde også erfaring med kommersiell BI. Den siste respondente var seniordirektør i et av Norges største BI-kompetansesenter. Han hadde kun erfaring med kommersiell BI, men hadde implementert løsninger for både SMB og store bedrifter. Dermed var alle nødvendige intervjuobjekt-statuser dekket.

- *Hawthorne-effekten*: I kvalitative intervju er ikke forskeren som holder intervjuet nøytral, men er interaktiv med egne tanker. Dette kan forstyrre intervjuobjektet, og forskeren kan derfor risikere å endre intervjuobjektets oppførsel. For å håndtere denne potensielle fallgraven, ble ledende spørsmål unngått, og forsker forsøkte å framstå med nøytrale holdninger til tema som ble diskutert.

5.3 Dataanalyse

Denne oppgaven hadde som utgangspunkt at Acando ønsket en vurdering på hvorvidt de kunne bruke en OSBI-plattform ved implementering av BI hos sine SMB-kunder. I starten hadde jeg derfor et bilde om at det var et litt enten/eller-forhold til bruk av OSBI, og at modenheten på plattformene ville avgjøre hvorvidt slike plattformer kunne brukes. Det ble derfor gjennomført et feltstudie for å gjøre en vurdering av en av de største OSBI-plattformene, med plattformens modenhet i fokus. Spørreundersøkelsen som ble gjennomført, hadde også høyt fokus på verktøyene og plattformene brukt under utviklingen. Etter hvert viste det seg at problematikk rundt implementering av BI hos SMB, ikke bare dreier seg om programvare, men også hvilke forutsetninger bedriften har, og hvordan bedriften går fram ved en evt. implementasjon. Problemstillingen endret seg derfor fra å omhandle hvorvidt OSBI er godt nok for SMB, til en problemstilling rundt optimal framgangsmåte for SMB-er som ønsker å implementere BI. OSBI ble da et undertema for SMB, hvor det vurderes hvilke typer bedrifter som kan benytte OSBI.

Det er tidligere gjort lite forskning med konkrete eksempler på BI-implementasjoner hos SMB-er. Den forskningen og teorien som er skevet på dette området, er stort sett på bransje-nivå, noe som fører til at dette studiet ikke har noe direkte sammenligningsgrunnlag når feltstudiet gjort på Acando skal diskuteres.

5.4 Refleksjon

For refleksjon over forskningsmetodene brukt, er Klein og Myers' evalueringssprinsipper [35] tatt i bruk. Disse prinsippene er ment å skulle gjøre forskere bevisste på hva som gikk bra og hva som gikk dårlig, samt vurdere validitet, reliabilitet og kvalitet på data som ble samlet.

5.4.1 Den hermeneutiske sirkelen

Den hermeneutiske sirkelen bygger på prinsippet om at all menneskelig forståelse oppnås ved å iterere over vurderinger av uavhengige deler, og den helheten som disse uavhengige delene former [35]. Sakt på en annen måte; ved å se på enkelte tilfeller for seg, og deretter sammenlignet med helheten, får man en komplett forståelse. I denne oppgaven var Acando et enkelt tilfelle, mens det norske BI-markedet kan sees på som helheten. Kunnskapen oppnådd gjennom feltstudiet hos Acando kunne brukes videre for å oppnå kunnskap om BI-markedet generelt, gjennom spørreundersøkelsen.

Ved å se på Acandos tilfelle, fikk jeg satt meg inn i BI som IT-område, og bli kjent med prinsipper, termer og vanlige utfordringer. Jeg fikk god innsikt i organisasjonen, og et intimt kjennskap til dilemma, frustrasjoner, rutiner, forhold og risikoer som typisk er en del av hverdagslivet til utviklere på BI-prosjekteter hos SMB-er. I ettetid leste jeg flere artikler om hva som oppfattes som de største utfordringene for en BI-implementasjon hos en SMB. Det ble da på flere områder klart at Acando skilte seg ut. De hadde for eksempel ikke en like stor risiko for implementasjonen som bedrifter som regel har. Dette skyldes at de brukte gratis programvare, og gjennomførte første del av utviklingen som et studentprosjekt, noe som førte til at kostnadene ikke ble like store som de ville blitt ved en tradisjonell BI-implementasjon. Dermed ble ikke korrekt utregning av ROI like kritisk. Acando skilte seg også ut ved at ingen av utviklerne hadde kunnskap om BI fra før. Dette førte til at det ble et stort "læringspreg" over hele utviklingsprosessen. Ellers er Acandos Trondheimskontor på grensen til det som regnes som mikro-bedrift, da de har ca 50 ansatte. Dette er 200 mindre ansatte enn hva som er øvre grense for å regnes som SMB². Dette kan ha ført til at det i mindre grad var problematikk relatert til de ansatte.

Ved de tre intervjuene som ble holdt, ble flere holdninger til OSBI dekket. De to som hadde jobbet med OSBI hadde andre holdninger til OSBI enn han som ikke hadde den samme erfaringen. Det var svært nyttig og interessant å få ulike synspunkt på samme tema, da jeg kanskje lettere greide å holde et kritisk syn på det som ble sakt.

5.4.2 Kontekstualisering

Prinsippet om kontekstualisering går ut på å gjøre en kritisk refleksjon over bakgrunnen til konteksten hvor forskningen ble gjort. Man får da se hva

²Totalt har Acando langt flere ansatte, men denne implementasjonen ble kun gjort for Trondheimskontoret, som har om lag 50 ansatte.

som var spesielt med denne konteksten, og se hvordan og hvorfor denne utviklet seg som den gjorde. Det er derfor ønskelig å reflektere over tankene og føringene Acando gjorde for sin BI-implementasjon, før jeg kom inn i bildet som utvikler og forsker.

Acando er en bedrift med kultur for å benytte OS-verktøy i utviklingen de gjør. De har mange års erfaring med kommersiell BI, og har med tiden blitt mer interesserte i å undersøke OSBI-alternativer, da de anså OSBI som noe de kanskje kunne tilby sine kunder. Acando leste litt om to av de største OSBI-plattformene, og bestemte seg for å gjøre en intern BI-implementasjon basert på en av disse. Acando lagde en kravspesifikasjon over det de ønsket å få utviklet, og lagde et studentprosjekt ut av dette. Studentene som gjorde utviklingen fikk lite hjelp under veis, da ingen ansatte hos Acando kjente til OSBI-plattformen fra før. Da jeg begynte å jobbe for Acando, som en av to sommeransatte, skulle det jobbes videre på det arbeidet som ble gjort på studentprosjektet. Utviklingen som var gjort tidligere var ikke gjort etter grunnleggende BI-prinsipper, og det meste av arbeidet måtte derfor gjøres om igjen. Dette førte til at jeg og min kollega brukte mesteparten av sommerjobben på å “rette opp” noe som allerede var gjort, men gjort feil. Dette kan ha gitt meg et feilaktig bilde av hvordan et tradisjonelt BI-prosjekt ser ut.

Acandos generelt positive holdning til OS-programvare kan ha påvirket mitt eget syn på OS, da jeg med tiden ble svært integrert i Acando. Dette kan ha blitt forsterket av det faktum at jeg kun har fått erfaring med en OSBI-plattform, og ingen erfaring med kommersielle BI-plattformer. Dette ble jeg bevisst på da jeg fikk se resultatene fra spørreundersøkelsen, og skjønnte at OSBI var langt mindre kjent enn hva jeg hadde forventet. Jeg innså da at jeg hadde startet dette studiet med et feilaktig bilde over OSBI's posisjon på BI-markedet.

5.4.3 Interaksjon mellom forsker og subjekter

Prinsippet om interaksjon mellom forsker og subjekt, går ut på at forsker må ha et kritisk blikk på hvordan data samles gjennom sosial interaksjon.

Under feltstudiet ble det samlet data gjennom deltakende observasjon. Jeg var en av to sommeransatte, og jeg snakket aldri til min kollega om graden av observasjon som ble gjort. Kollega visste at jeg skulle gjøre en vurdering av programvaren, men det ble aldri nevnt hvor detaljerte notater jeg tok av hele implementasjonsprosessen. Målet med dette var å usynliggjøre min rolle som forsker, for å få en mest murlig naturlig setting for observasjonen.

Min rolle hos Acando har vært noe udefinerbar. Jeg begynte som en sommeransatt student, som samtidig skulle gjøre en vurdering av programvaren som ble brukt. Etter endt sommerjobb ble jeg værende hos Acando, for å skrive masteroppgave for bedriften. Dette har ført til at jeg har blitt stadig mer integrert i bedriften, og kanskje har blitt opplevd som “en del av tappetet” for de ansatte. Det er mulig at jeg derfor fikk en mer “usensurert” sannhet om bedriften, samtidig som jeg selv kan ha mistet distansefølelsen overfor bedriften. Jeg tror at dette kan ha svekket min evne til å se problemer på andre måter enn bedriften, noe som har skapt et slags “tunnelsyn”. Et eksempel på dette er mitt forhold til OSBI. Jeg fikk i oppgave å vurdere om en OSBI-plattform var et godt alternativ for SMB-er, og tenkte aldri over at det kunne være andre faktorer enn programvarens modenhet som spilte inn på hvorvidt en SMB ville gjøre en god implementasjon eller ikke. Det ble i løpet av studiet klart at jeg også burde se på andre faktorer.

5.4.4 Abstraksjon og generalisering

Prinsippet om abstraksjon og generalisering går ut på å relaterer funnene til generelle konsepter.

Kunnskapen om BI som jeg tilegnet meg 1la. feltstudiet, var i stor grad generell BI-kompetanse, med BI-prinsipper, muligheter og utfordringer innenfor BI. Alle BI-implementasjoner følger de samme grunnprinsippene, med et datavarehus i grunnen, ETL steget for integrasjon, og deretter BI-funksjonalitet på toppen. Den innsikten jeg fikk i BI som felt, kan derfor sies å være generell for hele BI-markedet. Spørreundersøkelsen ga resultater som indikerte generelle trender for store bedrifter, for SMB-er, for bedrifter som benyttet kommersiell BI, og bedrifter som benyttet OSBI.

5.4.5 Flertolkninger

Prinsippet om flertolkning, sier at forsker må vurdere om det eksisterer mulige ulikheter i tolkninger hos deltakere som deltok i forskningen.

Da resultatene fra spørreundersøkelsen skulle analyseres, innså jeg at termen “data mining” er en term som brukes om forskjellige ting. Ett av intervjuobjektene gjorde meg oppmerksom på at noen bruker denne termen om vanlig drill-down analyse, mens termen i utgangspunktet oppstod som en betegnelse om mønstergjenkjenning ved hjelp av satte algoritmer. Denne mulige inkonsistensen i bruken av termen, har svekket reliabiliteten av dataen som omhandlet data mining. Dette har jeg forsøkt å ta hensyn til i analysen av resultatene.

Andre flertolkninger kan ha blitt gjort, basert på de ulike gruppene jeg har vært i kontakt med. Både ila. spørreundersøkelsen og under intervjuene, var det mange lederroller som besvarte spørsmål. Dette kan ha ført til at de “vanlige ansattes” meninger kanskje ble skygget over. Forståelsen av BI er i stor grad avhengig av hvem man snakker med, og det er en fare for at ikke alle disse gruppene er dekket godt nok i denne forskningen.

5.4.6 Misstanker

Prinsippet om misstanker krever at forsker vurderer sensitivitet til mulige “biaser” og systematiske skjevheter i data som ble samlet fra deltakere.

Den første misstanken som rettes, er mot meg selv. Myers [49] mener at en av styrkene ved etnografisk forskning og observasjon er at man kan se hva folk *gjør*, og samtidig se hva folk *sier de gjør*. Når jeg selv var deltaker i BI-utviklingen i feltstudiet, ble det vanskelig å observere “meg selv”. Det ble da kanskje problematisk å avdekke irrasjonaliteter og blindhet. Min rolle kan derfor ha svekket mulighetene for å samle korrekt data, samt å gjøre en god analyse av resultatene.

Jeg var helt fra starten av svært positivt innstilt til å benytte en OSBI-programvare i feltstudiet. Min hovedoppgave i dette studiet har vært å gjøre en vurdering av OSBI’s modenhet, slik at Acando kan tilby sine kunder OSBI som alternativ, dersom OSBI viser seg å være modent nok. Dersom min vurdering av OSBI’s modenhet antydte at det ikke er modent nok for SMB, vil ikke Acando få det resultatet som de har håpet på. Kanskje har jeg derfor hatt et underbevisst ønske om at OSBI-plattformen skulle vise seg å være moden nok for SMB-er. Dersom en forsker håper på et gitt resultat, kan dette føre til at forskeren velger å tolke resultater i den retningen som støtter den konklusjonen han hadde håpet på. Slike tilfeller vil svekke validiteten av konklusjonen som gjøres av forsker. Under veis i studiet ble jeg oppmerksom på at dette kunne skje, og forsøkte derfor å bli mer bevisst på de tolkningene jeg gjorde av resultatene.

Intervjuet som ble holdt av oppdragsgiver, i etterkant av BI-implementasjonen, var preget av rollekonflikt fra min side. jeg hadde tidligere jobbet på prosjektet selv, og var nå masterstudent som skrev masteroppgave for bedriften. Jeg var, som forsker, ute etter ærlige svar om hva som hadde gått bra, og hva som hadde gått dårlig, under implementasjonen. Jeg fikk kun positive svar relatert til utviklingsprosessen, og frykter derfor at intervjuobjektet ikke ville nevne negative sider av utviklingen, da jeg var en av to utviklere på prosjektet.

I det ene intervjuet som ble holdt i etterkant av spørreundersøkelsen, fikk jeg følelsen av at intervjuobjektet var forstyrret av Hawthorne-effekten³. Intervjuobjektet var nølende til et spørsmål, og spurte hva jeg ville gjøre dersom jeg ikke fikk det svaret jeg håpet på. Jeg fikk inntrykk av at han ønsket å svare det han trodde at jeg håpet at han ville svare. Det ble da presisert at jeg forholdt meg nøytral til konklusjonen som ville gjøres til slutt, og at enhver konklusjon var en like god konklusjon.

³Intervjuobjekt fikk påvirket atferd av forskerens tilstedeværelse.

Kapittel 6

Feltstudie: Acandos BI-løsning

Feltstudiet ble utført ved å se på et eksempel på bruk av OSBI hos en SMB i praksis. Målet med feltstudiet var å få et innblikk i BI som teknologi, SMB's bruk av BI, og hvor godt en OSBI-plattform fungerer for utvikling av en slik implementasjon i praksis.

Feltstudiet er utført på Acando og deres BI-løsning, basert på OSBI-plattformen Pentaho BI Suite. Implementasjonen ble gjennomført som et todelt studentprosjekt; først som et NTNU-prosjekt, og videre som en sommerjobb for to studenter (meg selv inkludert). Dette feltstudiet ble gjort som en del av sommerjobben. I første del av studiet ble case-studie-metodikken brukt, for å finne informasjon om utviklingen som var gjort under NTNU-prosjektet. Deretter ble det gjennomført en fokusgruppe med administrasjonen hos Acando, for å finne ut hva de ønsket å få ut av BI-implementasjonen, og hva de mente om arbeidet som var gjort så langt. Resten av feltstudiet bestod av deltakende observasjon på implementasjonen, samt noen intervjuer i etterkant av utviklingen.

Aktører i dette feltstudiet vil omtales på følgende måte:

- *Kundestyrt Prosjekt*: NTNU- prosjektet som bestod av 5 studenter, som gjennomførte faget TDT4290- Kundestyrt Prosjekt, høsten 2009
- *Sommeransatte*: De to studentene som jobbet for Acando sommeren 2010, meg selv inkludert
- *Veileder X*: Veileder for Kundestyrt Prosjekt
- *Veileder Y*: Veileder og sjef for de sommeransatte

- *Kollega X*: Min kollega på sommerjobben
- *Administrasjonen*: De ansatte hos Acando, som jobber med administrativt arbeid (rapportering)

Dette kapitlet er delt inn i fire seksjoner. Første seksjon beskriver Acando som virksomhet og Acandos behov for en BI-løsning. Seksjon 2 introduserer OSBI-plattformen, Pentaho, som ble brukt under implementasjonen. Tredje seksjon beskriver implementasjonsprosessen, som en kombinasjon av erfaringene gjort av Kundestyrte Prosjekt, og meg og min kollega. Fjerde seksjon presenterer de ulike aktørenes erfaringer og refleksjoner i etterkant av prosjektet. Femte og siste seksjon inneholder en helhetlig vurdering av prosjektet. Vurderingen baseres på utviklernes erfaringer gjort i implementasjonsprosessen, notater fra case-studiet, og intervjuer av ansatte hos Acando (både oppdragsgiver og sluttbrukere).

6.1 Bakgrunn

6.1.1 Acando

Acando er et Nordeuropeisk IT-konsulentselskap som ble startet i 1999 [2]. De tilbyr på den tiden generelle IT-løsninger, og fikk raskt økt kompetanse på ulike IT-områder. I 2003 slo de seg sammen med Frontec, et Stockholmselskap som ble opprettet i 1981. Acando hadde fram til da kun vært aktive i Sverige, men i 2006 slo de seg sammen med Resco, og ble da også aktive i Tyskland og Finland. I 2006 og 2007 opprettet Acando kontorer i Danmark og Storbritannia. De kjøpte da også opp det Norske selskapet Abeo AS, med ca 90 ansatte i Oslo og Trondheim. Abeo AS var resultatet av en fusjon i 2004 mellom Consult IT og eScienza, og ble deretter slått sammen med Addcom Innovation i 2005.

Acando har i dag i overkant av 1100 ansatte. De har kontorer i Norge, Sverige, Danmark, Finland, Storbritannia, og Tyskland. Konsernet omsetter for om lag en milliard NOK årlig, og har ca 30 000 aksjeeiere av konsernet. Acando er partner med programvareleverandører som SAP, Microsoft, Oracle og Sun Microsystems. Acandos visjon går ut på å utføre virksomhetsforbedringer ved hjelp av informasjonsteknologi, med effektiv prosjektgjennomføring og lave total kostnader. Leveranseområder er blant annet Arkitektur, Business Intelligence-løsninger, Bedrifts-systemer, Information Management, og Applikasjons-forvaltning. Acando har kunder både i offentlig og privat sektor, med en fordeling på ca 20 prosent i offentlig sektor, og 80 prosent i privat sektor. I den offentlige sektoren har de blant annet bidratt

mye til helse- og sentralforvaltning. I den private sektoren har Acando jobbet mest mot bransjene Telecom, Energi, og Bank og Finans.

Acando Norge har ca 100 ansatte, fordelt på kontorer i Oslo og Trondheim. Hovedkontoret ligger på Sluppen i Trondheim. På Trondheimskontoret er det om lag 50 ansatte, og det er på dette kontoret dette feltstudiet ble utført.

6.1.1.1 Steget mot BI

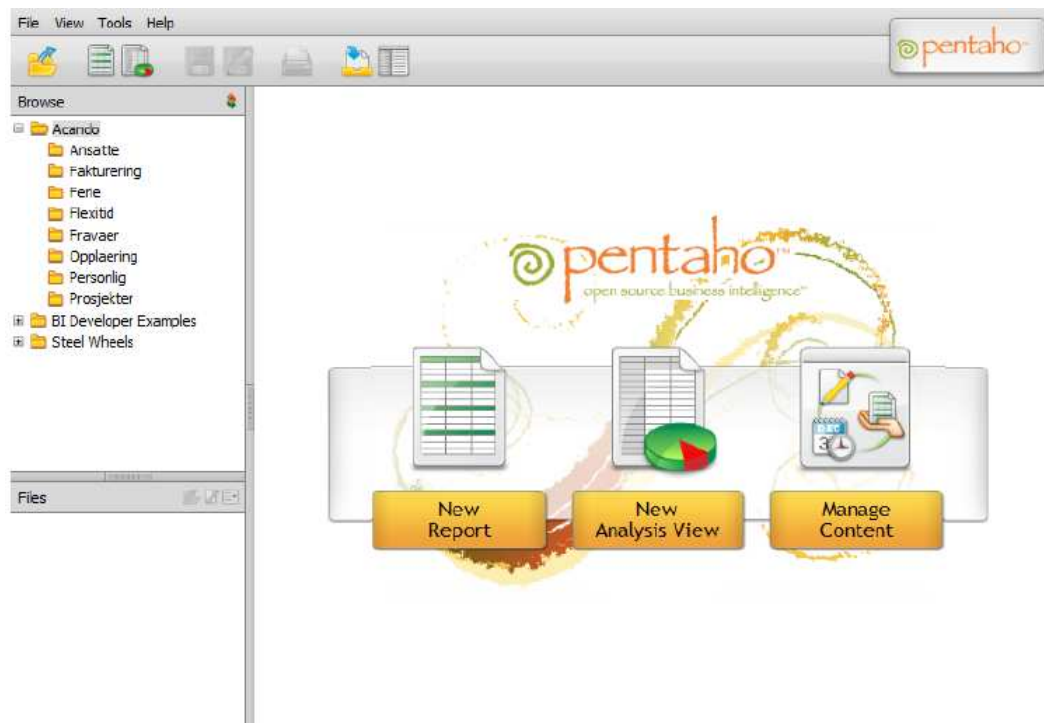
Administrasjonen hos Acandos Trondheimskontor (heretter kun omtalt som Acando) bruker i dag mye tid på rapportering relatert til regnskap, økonomi, ansatte og prosjekter. Mange av disse rapportene består av data fra ulike systemer, som må sammenlignes og samkjøres. Acando ser forbedringspotensiale rundt informasjonsflyten på flere nivå; samkjøring av relatert data, og enklere levering av data/informasjon til de ansatte. Datakilder som holder relatert data, mangler integrasjon, noe som gjør at sammenligning og rapportering av disse dataene må gjøres manuelt. Dette gjøres av spesielt kompetente ansatte, og “vanlige” ansatte har ikke tilgang til disse dataene, og er derfor nødt til å gå igjennom mellomledd for å få den dataen/informasjonen som de har behov for. Mange av rapportene er tidskrevende å lage, og må lages ofte. Det har derfor oppstått et ønske om å gjøre ressursbesparelser, ved å fjerne unødvendige manuelle prosesser. Acando har noe intern kompetanse på Business Intelligence, og begynte derfor å se på muligheten for å implementere en BI-løsning. Med denne ønsket de å integrere systemene brukt av Acando, samt automatisere de manuelle rapportene. Ved å gjennomføre implementasjonen som et studentprosjekt, ville de slippe for store kostnader rundt utviklingen. Utrekning av ROI ble derfor ikke veldig kritisk for Acando, da de i første omgang kun fikk kostnader for tiden brukt på å veilede studentprosjektet.

Acando har i flere år gjort BI-implementasjoner for sine kunder. Disse har vært basert på kommersielle løsninger, og Acando har hatt et ønske om å se på hvilke alternative OSBI-plattformer som finnes. Dersom det viser seg at en av de største OSBI-plattformene er et godt alternativ, kan Acando da tilby denne som mulig BI-plattform for sine kunder. Acando bestemte seg derfor for å gjøre sin egen BI-implementasjon basert på OSBI. Etter å ha undersøkt hvilke alternativer som fantes av OSBI sto valget mellom Pentaho BI Suite og Jaspersoft. Valget falt på Pentaho BI Suite, som et tilfeldig valg av de to.

6.2 Pentaho BI Suite

Pentaho-prosjektet ble startet i 2004 og har hovedkontor i Orlando i Florida, USA. Det fantes på dette tidspunktet flere OSBI-prosjekter, men dette var det første kommersielle OSBI-prosjektet på markedet. Pentaho BI Suite er ikke ett enkelt produkt- men en plattform; en samling programmer som jobber sammen for å bygge BI-løsninger [8]. Plattformen finnes i to versjoner; den kommersielle avgiftsbelagte versjonen *Pentaho Enterprise Edition* og den kostnadsfrie Open Source-versjonen *Pentaho Free Community Edition*. Pentaho har fem overordnede funksjonsområder:

- Integrasjon: Samling av data fra flere kilder til en felles database/datavarehus
- Rapportering: Brukere kan designe rapporter, lagre disse på Pentaho-serveren, og distribuere de automatisk. Ad-hoc-rapporter kan også lages, ved hjelp av en egen editor. Ad-hoc-rapporter kan aksesseres og redigeres av “vanlige brukere” gjennom web-browseren
- Analysering: Brukere kan analysere datagrunnlaget ved å gjøre drill-down's i dataen, gjennom web-browseren
- Data Mining: Mønster-gjenkjenning i datagrunnlaget ved hjelp av et eget verktøy
- Dashboards: Interaktive rapporter som gir output, enten basert på input fra bruker, eller basert på data i datagrunnlaget. Dashboards kan benyttes for å gi bruker beskjed dersom en hendelse inntreffer, for eksempel vha. en varselampe på skrivebordet

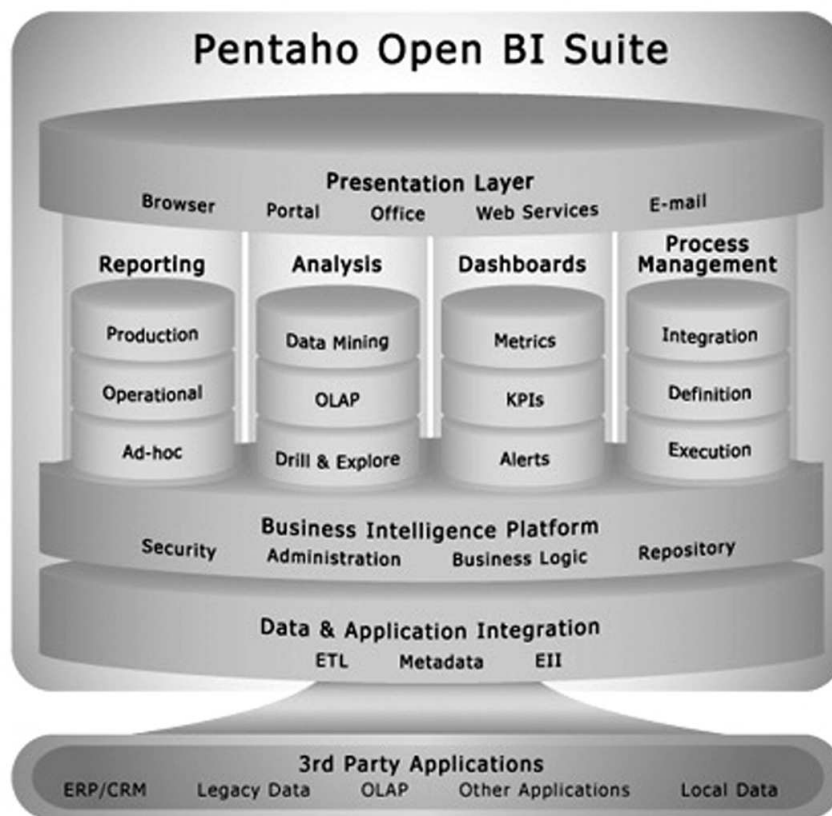


Figur 6.1: Grensesnittet i web-browseren, som gjør det mulig å lage rapporter og drill-down analyser.

Noen av komponentene som tilbys av Pentaho gjør helt enkle operasjoner, som autentisering og databasetilkoblingshåndtering. Andre komponenter leverer mer kompleks funksjonalitet, som visualisering av data, ved hjelp av grafer og diagrammer. Komponenter på høyere nivå er ofte avhengig av komponentene på de lavere nivåene, og det er derfor vanlig å se på verktøysamlingen til Pentaho som en stack, hvor høyere nivå tilbyr funksjonalitet nærmere knyttet til brukeren. Man kan dele Pentaho-arkitekturen inn i tre lag; integrasjonslaget, prosesseringslaget og presentasjonslaget.

Integrasjonslaget samler data fra en/ flere kilder (både interne og eksterne) og lagrer disse i et delt datavarehus/database. Det midterste laget, *prosesseringslaget*, også kalt Business Intelligence-laget, utgjør hovedfunksjonaliteten til BI-stacken; rapportering, analysing, dashboards og prosesshåndtering. Brukeren har ingen interaksjon med dette laget, da dette utfører ren prosessering av ulik form. Nærmest brukeren ligger *presentasjonslaget*, og her finnes verktøyene som kommuniserer med brukeren, igjennom for eksempel

web-browsersen eller rapportdesigneren.



Figur 6.2: En lagvis inndeling av Pentaho-stacken; Nedenifra og opp vises eksterne input-kilder, deretter *integrasjonslaget* (ETL), *prosesseringslaget* (BI-motorer) og øverst *presentasjonslaget*.

BI-serveren til Pentaho-plattformen inneholder et filsystem på toppen, som resten av applikasjonene bruker for å dele filer på serveren. Det er fullt mulig å bytte ut noen av programmene i Pentaho-stacken med andre programmer som gjør den samme jobben, så lenge de deler de samme standardene og filformatene. Pentaho og Jaspersoft deler noen programmer i sine BI-løsninger.

6.2.0.2 Pentaho-versjonene

De to versjonene er basert på de samme verktøyene og kan brukes på tvers av hverandre. Den kommersielle versjonen har noe funksjonalitet som ikke inkluderes i Open Source-versjonen; Ytelsesovervåking, dashboard-designer, større gard av sikkerhet, og automatisk backup-funksjonalitet. Dersom man kjøper Enterprise-utgaven får man også tilgang på en kunnskapsdatabase med dokumentasjon som ikke er fritt distribuert, i tillegg til support fra Pentahos ansatte. Support skjer over e-post og telefon.

6.2.0.3 Eksempel på bruk av Pentaho BI Suite

Flere Norske bedrifter og organisasjoner har valgt å benytte seg av Pentaho. Blant disse finnes en av Norges største TV-kanaler, TV2 [55]. TV2 benytter seg av Sumo Video Platform, som er en tjeneste som tilbyr streaming av TV-programmer over internett. TV2 har laget et eget produkt, TV2 Sumo, basert på denne. TV2 Sumo er et anerkjent produkt internasjonalt, og benyttes også av den svenske TV-kanalen TV4, og den finske kanalen MTV3. De som benytter seg av TV2 Sumo-løsningen får også Pentaho tilgjengelig. Ved å koble Pentaho BI Suite opp mot TV2 Sumo kan man se på TV-seernes bruksmønstre, og gjøre analyser av disse, på en måte som tidligere var svært vanskelig. Ledende seniorutvikler hos TV2 Sumo, Glenn S. Pedersen, har uttalt følgende til Pentaho;

“Ved å benytte Pentaho får vi større innsikt i seermønstre og transaksjoner. Pentaho BI Suite Enterprise Edition tillater oss å lett aksessere, trekke ut, analysere og distribuere ny informasjon mye raskere. Vi kan også lage bedre rapporter for hele organisasjonen”.

En annen stor organisasjon som benytter seg av Pentaho BI Suite er den Nord-Amerikanske energileverandøren Power Costs, Inc (PCI) [54]. PCI bruker Pentahos Enterprise Edition for å la brukere aksessere og analysere sitt eget strømforbruk.

Andre kjente selskaper som benytter Pentaho:

- Mozilla
- MySQL
- SUN (OpenOffice.org)
- SpecSavers

6.3 Implementasjon

Acandos BI-prosjekt startet høsten 2009, av 5 studenter som gjennomførte faget Kundestyrt Prosjekt [25] ved NTNU. Prosjektet varte i et semester (13 uker), og utgjorde totalt 1560 arbeidstimer, fordelt på de 5 studentene. Deres hovedoppgave var å utvikle en rekke rapporter ut ifra en kravspesifikasjon gitt av Acando. Studentene hadde hverken erfaring med BI-verktøy, eller BI som IT-område fra før. Deres veileder, Veileder X, hadde heller ingen erfaring med BI, og studentene måtte derfor lære seg BI og Pentaho på egen hånd. Prosjektet ble gjennomført smidig, med Scrum-metodikken, og inneholdt totalt fire sprints [25]. Et sett rapporter skulle leveres etter hver sprint. Prosjektgruppa hadde møter med Veileder X etter hver sprint, for å se på hvordan siste sprint hadde gått, og for å planlegge den neste. Arbeidet gjort av prosjektgruppa hadde størst fokus på utvikling av rapporter. Disse skulle baseres på fiktive data, slik at Acando slapp å gi de tilgang til sine kilde-systemer. Det ble dermed ikke gjort noen integrasjon mot de systemene som dataen skulle komme fra.

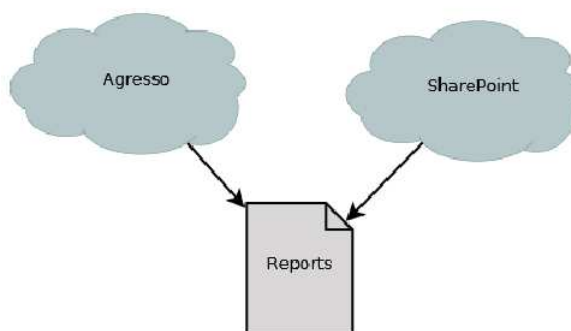
Allerede i første sprint av utviklingen fikk Kundestyrt Prosjekt en liste med rapporter som skulle lages. De måtte da rekke å sette seg inn i BI, modellere database, og gjennomføre ETL-steget, før de kunne begynne å lage rapportene. De fikk en stor feilestimering på denne sprinten, da de på forhånd hadde trodd at de kunne begynne med rapport-design med en gang. Databasemodelleringen og ETL-steget ble gjort så raskt som mulig, for å minimere feilestimeringen. De brukte XML-filer med fiktive data som testdata til utviklingen. Dette gjorde ETL-steget enklere, siden de da ikke trengte å lage et grensesnitt mellom Pentaho og kilde-systemene som skulle integreres. De resterende sprintene ble estimert mer korrekt, og de ble ferdige med de rapportene som Acando ønsket å få utviklet.

Jeg og min kollega overtok prosjektet et halvt år senere, sommeren 2010. Sommerjobben varte i 6 uker, med totalt 480 arbeidstimer. Vi hadde heller ingen kunnskap om BI fra før. Vår veileder, Veileder Y, hadde noe kunnskap om BI, og var derfor i stor grad inkludert i utviklingen. Målet med sommerjobben, var å fullføre implementasjonen som Kundestyrt Prosjekt hadde startet. Siden alle rapportene var laget, var det kun integrasjon mot Acandos systemer som gjenstod. Etter at sommerjobben var startet, bestemte Veileder Y at databasen som tidligere var laget, måtte lages om igjen. Databasen som Kundestyrt Prosjekt hadde laget, fulgte ikke BI-prinsippene om stjerneskjema og dimensjonalmodellering, noe som var ønskelig for optimal systemytelse. Jeg og min kollega brukte mye tid på modellering av ny database, og å gjennomføre ETL-stegene mot den nye databasen. Integrasjon mot Acan-

dos systemer ble ikke ferdig 1. august. sommerjobben. Det gjestet også arbeid på rapportene som var laget, da spørringer mot database måtte skrives på nytt for den nye databasen. Estimert tid nødvendig for fullføring av BI-implementasjonen ble satt til 554 timer[58].

6.3.1 Kravspesifikasjon for rapporter

Til å starte med ønsket Acando å fokusere på automatisering av de rapportene som tidligere ble laget manuelt. Disse rapportene omhandlet oversikt over økonomi, ansatt-data og prosjekt-data. Ved å dra nytte av primærfunksjonaliteten som BI tilbyr, ville de lage interaktive rapporter som endres når grunn-dataen endres, og distribueres til de ansatte, enten automatisk eller ved forespørsel. Datakildene som måtte integreres var Microsoft Sharepoint, som holdt prosjekt-data, og ERP-systemet Agresso, som holdt ansatt-data.



Figur 6.3: Acando ønsker å få rapporter basert på en kombinasjon av data fra Agresso og Sharepoint.

Nedenfor er noen eksempler på noen av rapportene som Acando ønsket å få utviklet.

Rapporter for de ansatte:

- Ferie til gode
- Planlagt ferie
- Fraværsrapport

Rapporter for ledelsen:

- Planlagt ferie for alle ansatte: Tabelloversikt, og graf over hvor mange som er på ferie samtidig
- Overtidsrapport: Oversikt over hvem som har jobbet for mye overtid i henhold til arbeidsmiljøloven
- Prosjekt-fakturering: Oversikt over hvor mye som er fakturert på de ulike prosjektene
- Ansatt-fakturering: Oversikt over hvor mye hver ansatt har fakturert i gitte tidsintervall, samt faktureringsgrad¹ for de ansatte
- Prosjekt-aktivitet: Oversikt over de 5 mest aktive prosjektene, med ulik tilhørende informasjon

Acando ønsket også å etter hvert ta i bruk ad-hoc-rapportering, for å gjøre det mulig for de ansatte å velge selv hvordan felt de skulle ha med i rapporten, uten for mye teknisk konfigurering. Administrasjonen hos Acando skulle da kunne lage halvferdig rapporter, som kunne redigeres av de ansatte, innenfor gitte rammer. På denne måten ville vanlige ansatte slippe å få opp alle attributter i datagrunnlaget som valg, men heller kun relevante alternativer. De ville også slippe å gjøre grunnarbeidet til selve rapporten, som omhandler databasetilkobling, SQL/MDX-spøringer og design.

6.3.2 Database

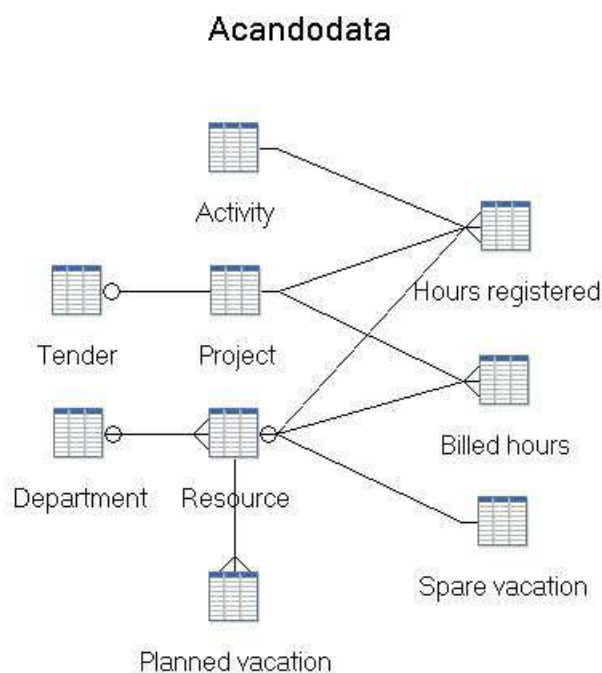
6.3.2.1 Databasemodell

Da Kundestyrt Prosjekt skulle lage database, hadde de ikke gode nok kunnskaper om BI til å vite hvordan dette burde gjøres. De brukte derfor den modellen som de kjente fra NTNU; ER-modellen. Dette er, som nevnt i Kap. 1, sjelden en god idé for en BI-løsning, men Kundestyrt Prosjekt ble ikke klar over dette før et stykke ute i prosjektet. I tillegg rapporterte de at de var usikre på hvordan Acandos komplette datagrunnlag så ut, noe som førte til at de måtte modellere databasen “i blinde”. For å gjøre det enkelt valgte de å kopiere alle feltene fra XML-filene rett inn i databasen. Databasemodellen fikk derfor en egen tabell for hver XML-fil.

¹Faktureringsgrad viser hvor mange timer en ansatt har fakturert for, i forhold til hvor mange timer han har jobbet.

Filnavn	Forklaring
Aktiviteter.xml	Oversikt over aktivitetene som er tatt ut fra Agresso, med navn og nr. Disse blir lagt inn i databasen med navn og ny primærnøkkel
Avdelinger.xml	Oversikt over avdelingene innad i Acando. Dette blir lagt inn som departments i den nye databasen
Ferie_tilgode.xml	Antall feriedager til gode fra tidligere år, registrert på ressurs
Planlagt_ferie.xml	Planlagt ferie som enten er "Godkjent", "ikke godkjent" eller "Ikke behandlet". Ferien er knyttet opp mot startdato og sluttdato. Startdato er første feriedag som er virkedag, og sluttdato er siste feriedag som er virkedag.
Prosjekter.xml	Prosjekter som er hos Acando, føres med prosjektnummer og navn
Ressurser_2.xml	Ansatte innad i Acando, med personlige data registrert slik som fødselsdato, ansettelsesdato osv.
Tilbud.xml	Prosjekter som er registrert som tilbud. Det vil si at administrasjonen i Acando regner tilbud på disse prosjektene
Timer.xml	Timeføring fra ressurser hos Acando. Registreres av en ressurs, på et prosjekt, med en aktivitet og år, uke og antall timer
Utfakturert.xml	Timer og beløp som skal faktureres. "tim" er antall fakturerbare timer ført av konsulent. "ant" er antall fakturerbare timer ligger klare til fakturering, disse timene kan redigeres av lønnsseksjonen i Acando. "tot" er beløpet som faktisk ble fakturert, i norske kr.

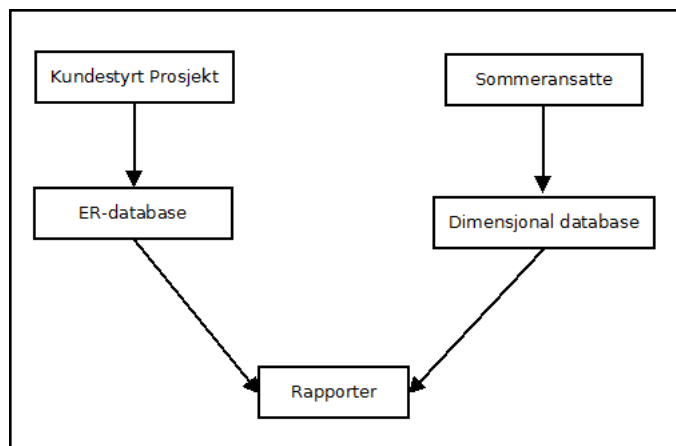
Figur 6.4: Beskrivelse av XML-filene som gav simulert/fiktiv test-data til implementasjonen.



Figur 6.5: Databasemodellen som ble utviklet av Kundestyrte Prosjekt.

Veileder Y mente at en ER-modell var svært ugunstig for Acandos langtidsperspektiv, da spørringer mot databasen etter hvert ville gå treigt med en ER-modell, etter hvert som datagrunnlaget vokste. I tillegg ville dataen bli uhåndterbar og vanskelig å finne fram i. Han mente i tillegg at det ville bli lettere å gjøre drill-down analyser, dersom databasen var modellert hierarkisk, slik som den er i den dimensjonale modellen. Jeg og min kollega fikk derfor i oppgave å gjøre modellering av BI-databasen om igjen, og da følge BI-prinsippene om stjerneskjema og dimensjonalmodellering.

Veileder Y hadde selv erfaring med BI, og ga oss en innføring i prinsippene om dimensjonalmodellering av stjerneskjema. Vi opplevde databasemodelleringen som en krevende oppgave, ettersom at vi kun hadde erfar-



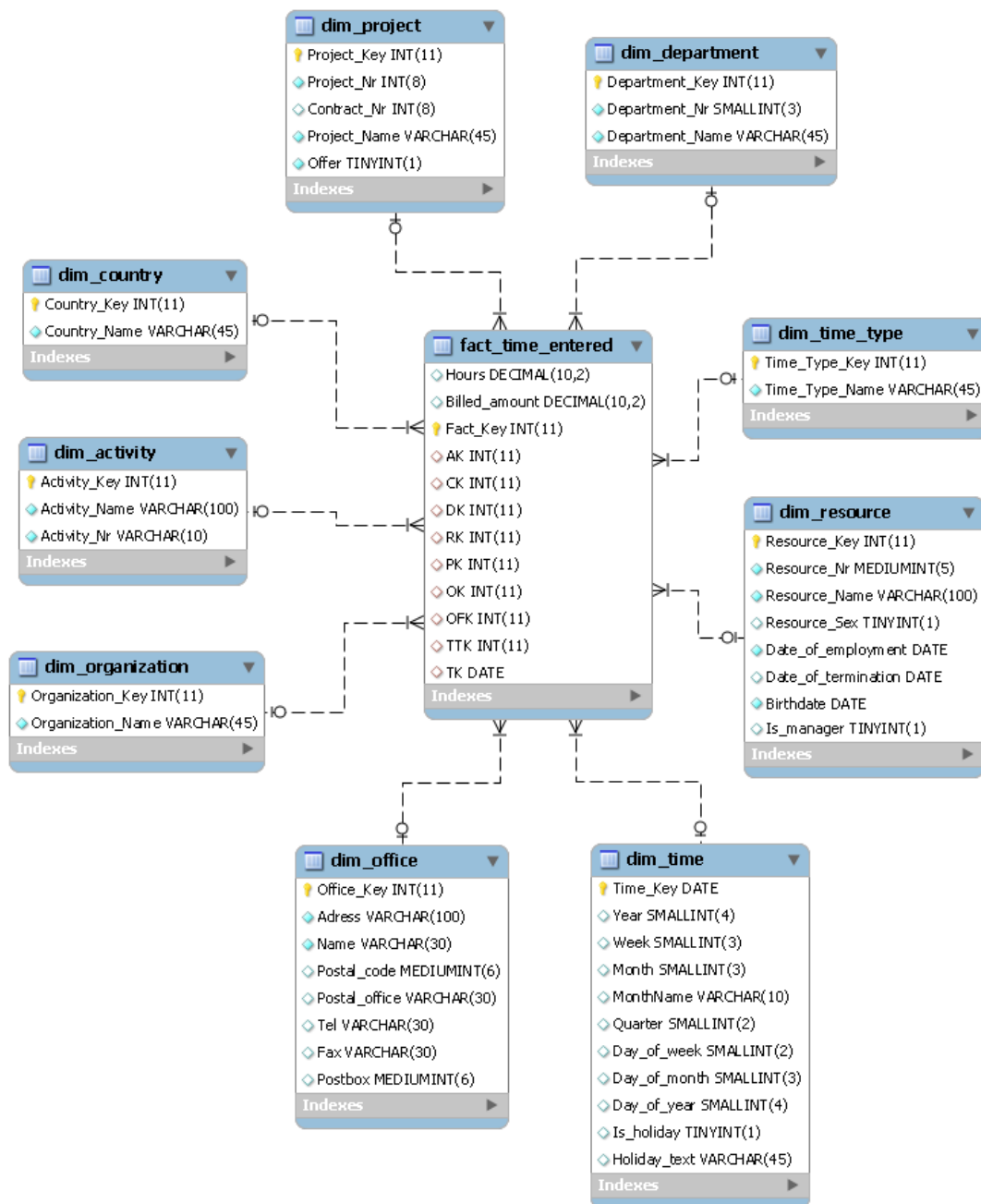
Figur 6.6: Figuren viser at det ble laget to forskjellige databaser, men de samme rapportene ble brukt.

ing med ER-modellering, som følger helt andre prinsipper. ER-modellering fokuserer på å bruke liten plass, mens dimensjonalmodellering ikke tar hensyn til plass, men heller at uthenting av data skal gå så raskt som mulig. Hierarkisk oppbygging av tabeller, og skille mellom faktatabeller og dimensjonstabeller var helt fremmed for oss. Vi måtte i stor grad støtte seg på Veileder Y, og først etter en ukes jobbing med databasemodellen, følte jeg at jeg mestret modellen som ble brukt.

Utdrag fra notater etter endt uke:

“Hadde det ikke vært for at Veileder Y pekte oss i rett retning hadde vi nok endt opp med å beholde databasen som Kundestyrt Prosjekt hadde laget. Dimensjonalmodellering er veldig annerledes enn ER-modellering, og prinsippene er vanskelig å skjønne i starten. Jeg er glad jeg hadde Kollega X å støtte meg på, og jeg tror vi hjalp hverandre å få en bedre forståelse.”

Resultatet ble stjerneskjemaet i figuren nedenfor.



Figur 6.7: Stjerneskjema av databasen til BI-løsningen, laget av meg og min kollega.

Stjerneskjemaet holder to målinger i faktatabellen; *Antall Timer*, og *Fakturert Sum*. Den førstnevnte ble fastslått med en gang, ettersom det er nettopp timer som føres av de ansatte, og det er disse som er av interesse ved analyser. Timene skal rapporteres i de fleste rapportene, i form av ulike registrerte timer, som for eksempel:

- Ferietimer (gjennomført og planlagt)
- Timer jobbet
- Fleksitid-timer tatt ut i avspasering
- Sykefraværstimer
- Timer brukt på opplæring

Dermed fikk faktatabellen navnet *fact_time_entered*, og dimensjonene som ble tilhørende² faktatabellen beskriver følgende:

- Tidspunkt for timene
- Person (ressurs) som registrerte timene
- Land, organisasjon, avdeling og kontor hvor timene ble ført
- Aktivitet og prosjekt som timene ble registrert på
- Tidstype

Den andre målingen som ble tatt med i faktatabellen var *Fakturert Sum* (Billed Amount), som holder data om en fakturering til kunde. Denne hadde tilknytning til nesten alle de samme dimensjonene som *Antall Timer*-feltet. Vi hadde valget mellom å lage to stjerneskjema; ett for hver måling. Vi måtte da vurdere hva vi skulle gjøre, og ønsket å gjøre det enklest for oss selv og Acando, samtidig som vi ønsket optimal ytelse for spørringer mot databasen. Det ville bli lettere for brukere å håndtere databasen med bare ett stjerneskjema, samtidig som vi ville spare oss selv for en del jobb ved å ha alt i ett. Veileder Y mente at ytelsen ikke ville bli merkverdig påvirket av å ha begge målingene samlet, fordi målingene så og si hadde de samme dimensjonene, med unntak av to. Vi bestemte oss derfor for å kun ha ett stjerneskjema. I ettertid så vi at vi sparte mange timers arbeid på dette valget, men vi fikk aldri sett om det påvirket ytelsen til systemet.

²Tilhørende betyr her at tabellene er knyttet ved hjelp av fremmednøkler.

6.3.2.2 Kube

På toppen av stjerneskjemaet bygde vi en OLAP-kube, som beskrevet i kap. 1. Kuben ble laget i verktøyet Schema Workbench, som er et verktøy utviklet i samarbeid mellom Pentaho og OSBI-konkurrenten Jaspersoft. Dette programmet ga et enkelt grafisk brukergrensesnitt, og var svært brukervennlig. I tillegg til å bygge kube, kunne man skrive MDX-spøringer mot kubene. Designet av den ene kubene tok mindre enn en dag. Det eneste problemet som oppstod under veis var at det ble fylt inn feil verdi på ett av feltene, noe som resulterte i en uforståelig og missledende feilmelding. Mange timer gikk bort til å finne ut hva som var feil, og det endte med at hele kubene måtte bygges om igjen. Det viste seg da at feltet `Month_name` i tids-dimensjonen var fylt inn med et tall i stede for en `String`. Vi ble overrasket og irriterte over at det i det hele tatt var mulig å skrive inn feil datatype, da det førte til at hele kubene ble ugyldig. Resultatet fra kubene ble lagret i et skjema i XML-format.

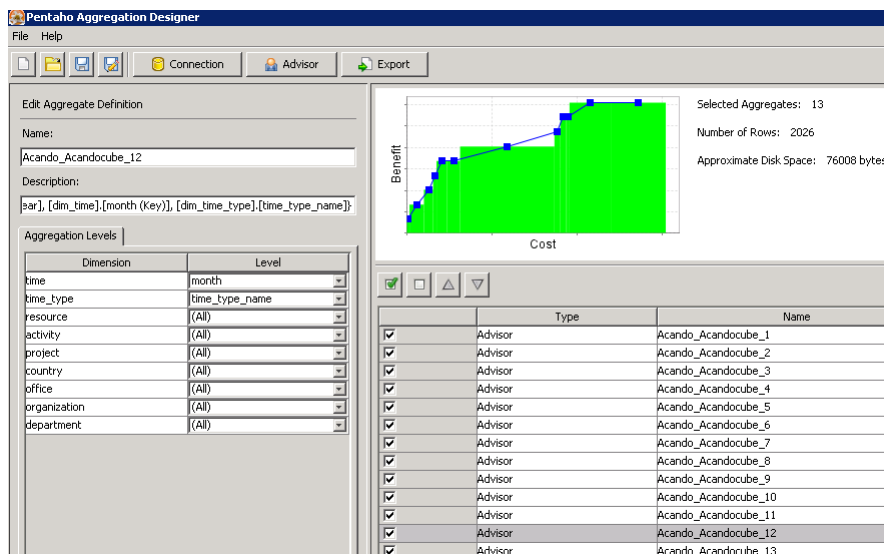
Vi hadde missforstått hvordan kubene fungerte som datastruktur, og trodde først at hvert stjerneskjema skulle ha en kube på toppen med de nøyaktig samme feltene. I ettertid skulle det vise seg at man kan ha ulike kuber som bruker ulike data fra stjerneskjemaet. Dette ble først klart etter fullført kube-utvikling, noe som resulterte i at det bare ble laget en kube, denne med all den samme dataen som finnes i stjerneskjemaet.

Helhetsinntrykket av kube-designeren var at den var noe umoden, men relativt grei å bruke. Feilmeldingene var varierende, da feilmeldinger for MDX-spøringer var svært informative, mens feilmeldingene for feil i selve kubedesignet var svært dårlige.

6.3.2.3 Aggregering av data

Ved å aggregere over datagrunnlaget vil noen verdier forhånds-kalkuleres og lagres, slik at man slipper å regne de ut for hver spørring. Slike forhåndskalkuleringer kalles aggregater, og Pentaho tilbyr et eget verktøy for å autogenerere slike aggregat-tabeller. Aggregering skulle vise seg å være en svært enkel prosess. Jeg og min kollega måtte først sette opp databasetilkobling, kobler til kube, og deretter kjøre en funksjon som kalles `Advisor`. Denne funksjonen regnet ut anbefalte aggregat-tabeller, ved å se på hvilke utregninger som gjøres oftest, og dermed er mest nyttig å forhåndskalkulere. Etter autogenerering av tabeller, lagrer Pentaho Aggregation Designer disse i databaseskjemaet. Det er en svært avansert algoritme som regner ut tabellene, og denne fungerer slik at den laster inn databaseskjemaet som ligger under kubene, og

analyserer dette for å se hvilke aggregat-tabeller som det lønner seg å lage. Algoritmen var for avansert til at vi greide å skjønne den, men i følge flere forum, skulle den være like god som tilsvarende verktøy for de store kommersielle konkurrentene.



Figur 6.8: Skjerm bilde av aggregat-designeren.

Vi opplevde verktøyet som svært enkelt å bruke, da hele aggregeringsprosessen tok mindre enn en time. Vi møtte noen problemer relatert til disse tabellene i ettertid av den første autogenereringen. Det ble konflikt når nye felt ble lagt til å databaseskjemaet, uten at vi skjønnte hvorfor. Det viste seg etter litt leting at aggregat-tabellene måtte oppdateres i designeren dersom nye database-felt ble lagt til, slik at de forhåndskalkulerte verdiene ble korrekte. Utdrag fra dagboken min:

“Da vi skjønnte hva som var feil syntes vi i grunn det var veldig intuitivt, men vi hadde likevel ikke greid å tenke oss til det selv. Det var dumt at det ikke stod noe notis om dette i dokumentasjonen til Aggregat-designeren. Vi fant løsningen ved å lete litt etter lignende problem på Pentaho-forumene.”

6.3.2.4 Databasetilkobling

Kundestyrte Prosjekt brukte mye tid på å sette opp databasetilkoblinger for hver rapport de laget [25]. Det viste seg etter hvert at det var mulig å ha

en universal tilkobling til databasen. Denne ble opprettet i Admin-console (gjennom Web-browsersen), og deretter ble rapportdesigneren konfigurert til å bruke denne tilkoblingen. Alle verktøyene som benyttet seg av data fra databasen måtte ha en slik database-konfigurasjon. Jeg og min kollega leste i rapporten til Kundestyrt Prosjekt at dette var mulig, og unngikk derfor å gjøre den samme feilen.

Et stort problem som vi opplevde under veis i utviklingen, var at dersom man endret noe i databasen eller kuben, ville ikke programmene som tok i bruk dataen oppdateres automatisk. Det ble etter hvert mye å holde styr på når 4-5 ulike verktøy tok i bruk det samme datagrunnlaget. For eksempel ble ikke kubene som var lastet inn i Pentaho Report Designer oppdatert dersom kubene på Serveren ble endret. Dette problemet ble vi oppmerksomme på først etter flere uker med utvikling.

6.3.3 Integrasjon

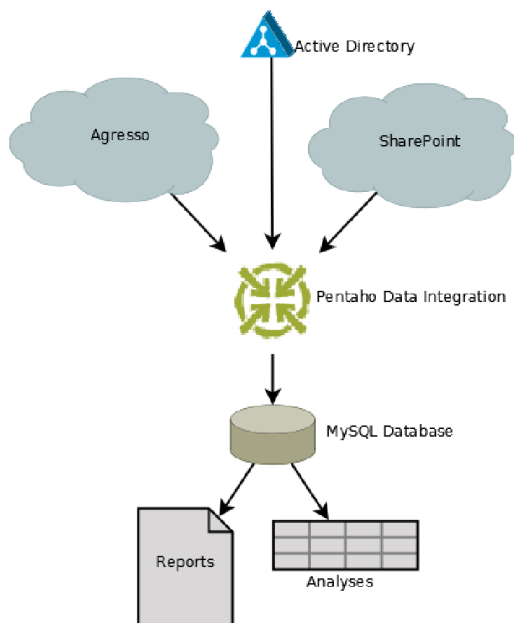
En av hovedfunksjonalitetene til BI er å integrere flere datakilder, både interne og eksterne. I Acandos tilfelle var det tre eksterne systemer som skulle integreres; Active Directory (AD), Agresso og Sharepoint. Det finnes to primærtyper for integrasjon; en for brukertilgang, og en for integrasjon av datakildene. AD tilhører den førstnevnte, mens Agresso og Sharepoint tilhører den sistnevnte.

Etter at datbasemodellen var komplett, kunne databasen fylles med data. Kundestyrt Prosjekt fylte, som nevnt, sin database med fiktive data, og gjorde derfor ingen integrasjon. Jeg og min kollega fullførte integrasjon mot AD, og begynte å integrere mot Agresso og Sharepoint.

6.3.3.1 Active Directory

AD er Acandos brukerkatalog for brukerhåndtering, brukerrettigheter og brukerautentisering. Jeg og min kollega skulle gjøre en komplett operasjonell integrasjon mellom AD og Pentaho, slik at Acandos ansatte ble registrert som brukere med gitte rettigheter i BI-implementasjonen. Vi måtte derfor sette oss inn i hvordan LDAP³ og AD fungerte sammen. I Enterprise-versjonen finnes et enkelt brukergrensesnitt hvor man setter opp koblingen mellom AD og Pentaho-serveren. En slik kobling måtte hardkodes inn i systemfilene (XML) til Pentaho Serveren, ettersom vi brukte den helt åpne versjonen av Pentaho, som ikke tilbydde slik funksjonalitet. Hvordan man gjorde dette fant vi delvis ut av ved å lete på Pentaho Community-forumene,

³LDAP er en protokoll for oppslag i brukerkataloger.



Figur 6.9: Acandos visjon om en ferdig integrasjon.

og delvis ved å dele erfaringer med en italiener som min kollega ble kjent med over chatte-tjenesten IRC. AD-integrasjonen var en tidskrevende oppgave, da det ikke fantes tilgjengelig dokumentasjon. Da vi til slutt fant ut hva som måtte gjøres, viste det seg at det bare var en enkelt setning som måtte skrives i den ene konfigurasjons-fila til BI-serveren. Integrasjonen mot AD var en frustrerende jobb, fordi vi visste at det kunne vært flere dager spart dersom vi hadde hatt bedre dokumentasjon tilgjengelig, eller dersom vi hadde brukt Pentaho Enterprise-versjonen, med nødvendig funksjonalitet innebygd.

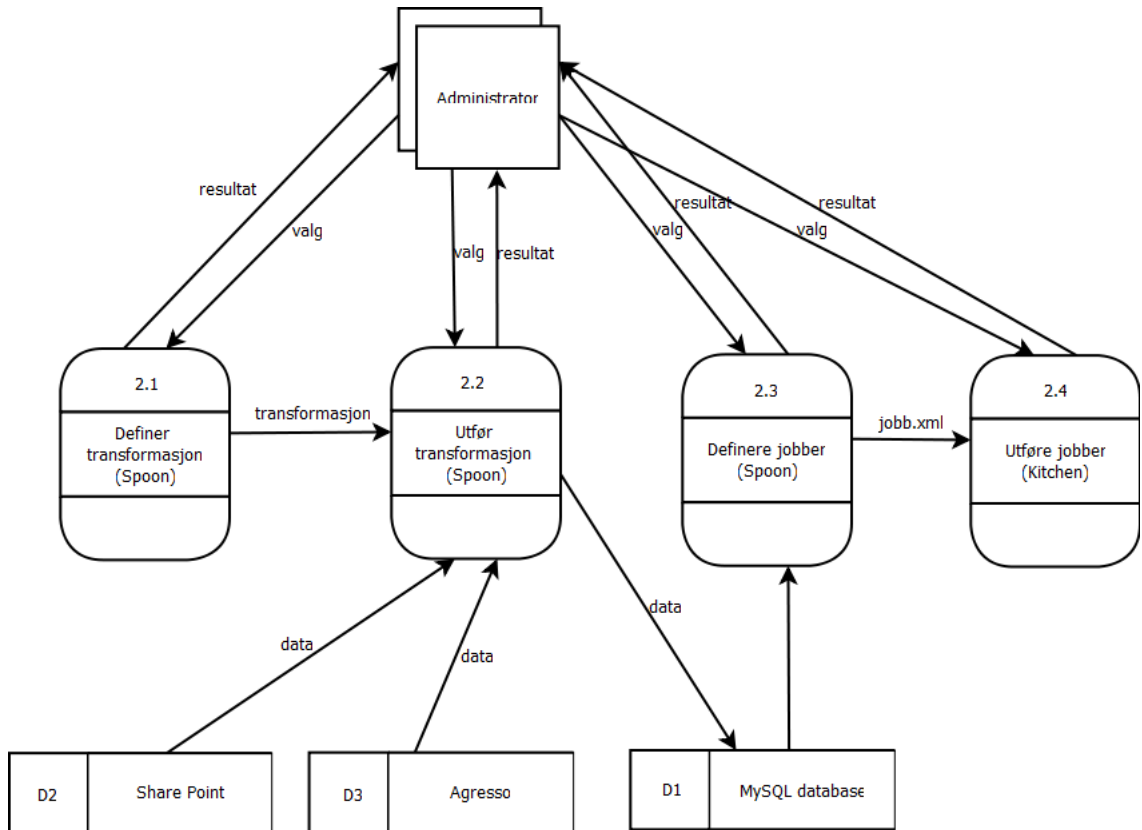
6.3.3.2 Sharepoint og Agresso

Sharepoint benyttes av Acando for at ansatte skal registrere data om prosjektene de jobber på, og data om planlagt ferie. I Agresso registrerer ansatte timer, og hver timeføring tilhører en aktivitetstype. De mest vanlige aktivitetene er Flexitid, Konsulentbistand, Egenopplæring, Sykdom, Rådgiving, Gjennomført ferie, og Pålagt overtid.

Data som kommer fra Agresso og Sharepoint lagres først i et XML-dokumenter, før disse sendes videre til Pentaho. XML-grensesnittet fungerer

dermed som en Wrapper, hvor BI-løsningen ikke trenger å ta hensyn til hvor dataen kommer fra. Man kunne tatt dataene direkte ut fra Agresso eller Sharepoint, uten å ha mellomstasjonen med XML-filer. Ved bruk av dette mellomleddet mellom Pentaho og støttesystemene standardiseres en måte å ta ut data fra systemer. ERP-systemene kan endres uten at jobbene og transformasjonene trenger å endres, så lenge det nye ERP-systemet retter seg etter det definerte XML-grensesnittet.

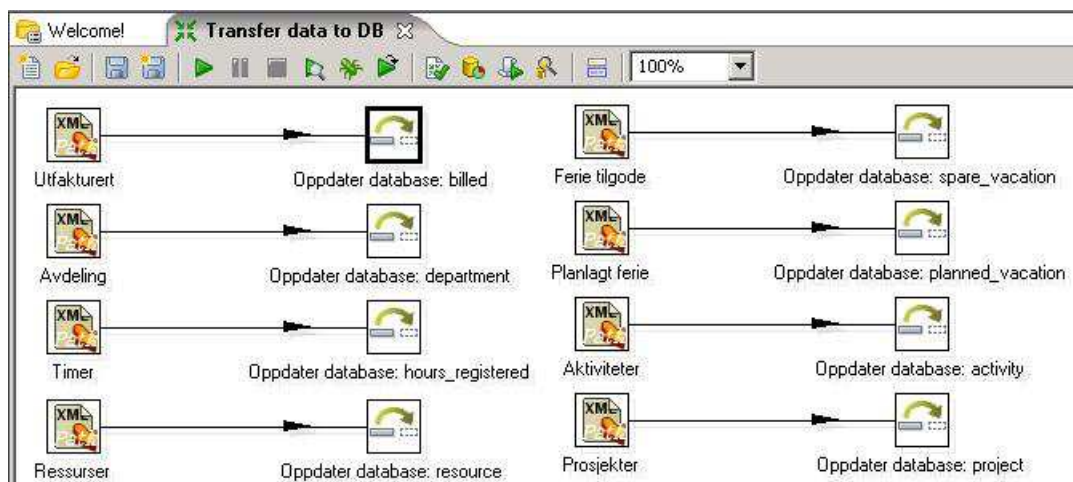
Integrasjon av datakilder består av å utføre ETL-stegene; hente ut data, gjøre om data til ønsket format, og laste opp omgjort data til datavarehus. Selve ETL-steget gjøres i Pentaho Data Integration, som består av flere verktøy, blant annet Spoon og Kitchen [53]. Vi fikk inntrykk av at Spoon er et svært populært Open Source-verktøy, som også i stor grad benyttes utenfor både Pentaho og BI-området, som et enkeltstående integrasjonsverktøy. Vi opplevde Spoon som svært brukervennlig, med et grafisk brukergrensesnitt for design av transformasjoner og jobber. Man definerer datakildene til input-data, og hver integrasjon kan ha n antall datakilder. De n datakildene kan være av n ulike typer, og her finnes svært mange alternativer (Excel, XML, enkel tekstfil, RSS, database, osv.). Deretter spesifiseres hvilke data som skal hentes ut, og hva som skal gjøres med denne dataen. Det kan for eksempel settes opp en join mellom data fra flere kilder. Det finnes mer enn 100 ulike typer transformasjoner i Spoon, men utviklere kan også lage disse selv vha. Javascript. Når man har fått dataen på det formatet man ønsket, sendes den videre til database/datavarehus. I figuren under vises et dataflyt-diagram over integrasjonsprosessen.



Figur 6.10: Dataflytdiagram for transformasjoner og jobber som lages i Spoon.

I flytdiagrammet er det laget fire del-prosesser for transformasjonene og jobbene. En jobb er en transformasjon som gjøres ved et fast tidsintervall, eller på kommando. Transformasjonen lages og kjøres i Spoon. Etter kjøring av transformasjonen vil integrasjonen være fullført. Dersom man ønsker at det skal være mulig å oppdatere databasen automatisk, lages jobber for å gjøre dette.

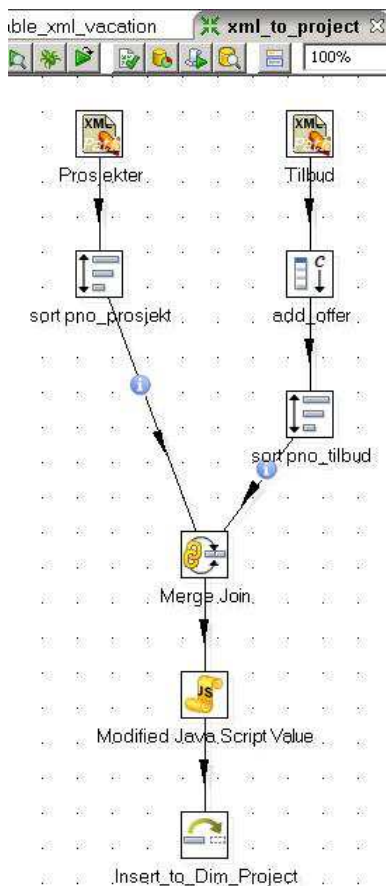
Kundestyrt Prosjekt lagde én tabell i ER-databasen, for hver XML-fil med fiktiv data. Ettersom de ikke hadde noen joins, ble det en enkel rett-ram-transformasjon, som vist i skjermbildet nedenfor. De lagde kun transformasjoner, siden datagrunnlaget aldri endret seg.



Figur 6.11: Skjermbilde fra transformasjonene som ble laget av Kundestyrt Prosjekt.

Jeg og min kollega tok i bruk de samme XML-filene under utviklingen, da det var bestemt at disse XML-filene skulle fungere som et grensesnitt mellom kildesystemene og Pentaho. Etersom hele databasen var bygget om, måtte også uthenting av data til databasen gjøres på nytt. Databasen bestod ikke lenger av tabeller direkte kopiert fra XML-filene, men tabeller som tok data fra flere ulike XML-filer, og flettet sammen disse. Det ble da nødvendig å joine og merge ulike kilder. I eksempelet nedenfor joines to XML-filer, som begge holder data om prosjekter. Resultatet blir ferdig utfylte Prosjekt-dimensjonstabeller i SQL-databasen. Det som skjer i selve transformasjonen er at feltene i Prosjekt.xml sammenlignes med feltene i Tilbud.xml for å se hvilke prosjekter som for øyeblikket er på anbud/tilbud. De prosjektene som er på anbud får et boolean-felt i prosjekt-tabellen satt til *true*. Selve joinen var enkel å gjennomføre fordi prosjektene i begge XML-filene var lagret med samme prosjekt-ID. Noen av dimensjons-tabellene ble laget i SQL-skript, mens andre ble laget ved hjelp av JavaScript i Spoon. For eksempel ble tidsdimensjonstabellen fylt ved hjelp av et SQL-skript, som gikk gjennom alle datoer *fra* en gitt dato *til* en gitt dato. Skriptet fylte tabellen med en oppføring for hver av disse dagene, med gitte attributter (eks. år, mnd, uke,

dag). Dette skriptet brukte vi flere dager på å skrive, og ca en uke etter at skriptet var ferdig, fant vi ut at det fantes en egen funksjon for å fylle tidsdimensjonen i verktøyet Spoon. Dette opplevde vi som svært ergelig, da det var flere dager med bortkastet arbeid.



Figur 6.12: Skjerm bilde av en av transformasjonene, som ble laget av min kollega.

Veileder Y var interessert i hvorvidt Pentaho gjorde det mulig å kun laste opp nye data i datavarehus. Ved å slippe å laste opp samme data flere ganger, vil man spare mye tid under opplastingen. Vi undersøkte dette, og fant ut at Pentaho støttet en slik inkrementell opplasting. Vi forsøkte ikke å teste det, siden vi bare brukte fiktive data, og dermed aldri fikk “nye data” for å kjøre en inkrementell opplasting. Vi dokumenterte heller hvordan dette skulle

gjøres, slik at de som evt. tok over utviklingen kunne implementere inkrementell opplastning. Pentaho data integration støtter også parallell kjøring av jobber, noe som vil redusere kjøretiden betraktelig. Dette ble heller ikke testet ut. Vi fikk inntrykk av at Pentaho har svært god ytelse for integrasjon, og at denne delen av Pentaho har et godt rykte på BI-markedet. Det eneste som manglet i den helt åpne versjonen, som fantes i Enterprise-versjonen, var prosessovervåking, som vil kunne forbedre ytelsen ytterligere.

6.3.4 Rapportering

Pentaho har støtte for to typer rapporter; “vanlige” rapporter og ad-hoc-rapporter. Vanlige rapporter lages fra grunnen av i Pentaho Report Designer, og krever en viss teknisk kompetanse. Ad-hoc rapporter lages i Pentaho Metadata Editor, og User Consol. Denne gjør det mulig for ikke-teknisk personell å lage sine egne rapporter, basert på halvferdige rapporter. Acando valgte å fokusere på de vanlige rapportene, ettersom de fleste rapportene som skulle lages skulle genereres automatisk, uten noen brukerinteraksjon på design-fronten. Ad-hoc rapportering ble også testet ut.

Hovedjobben til Kundestyrt Prosjekt var å lage et sett med rapporter i Pentaho Report Designer. Acando satte da krav til hvordan data rapportene skulle inneholde, og hvordan de skulle se ut. For hver av de fire sprintene i prosjektet, lagde Acando en ny liste over ønskede rapporter. Når Acando så resultatene etter hver sprint, kunne de basere de nye rapport-forespørsler på mulighetene som etter hvert ble synlige. Designet av alle rapportene som Kundestyrt Prosjekt hadde laget, kunne brukes videre av meg av min kollega, men datainnholdet måtte hentes inn på nytt.

6.3.4.1 Fokusgruppe

I starten av sommerjobben gjennomførte vi en fokusgruppe med Administrasjonen, som var ansvarlige for de manuelle rapportene. Dette ble gjort for å få tilbakemeldinger på rapportene som var laget av Kundestyrt Prosjekt, og for å få et innblikk i hva Administrasjonen så for seg å få ut av BI-implementasjonen. Fokusgruppen ble gjennomført som en åpen diskusjon rundt rapportene, som Administrasjonen hadde fått sett på forhånd. Det kom ila. fokusgruppen fram ting som burde endres, så vel som ting som var bra. De fleste kommentarene gikk på utseende av rapportene, men de hadde også tilbakemeldinger som avdekket feiltolkninger av datagrunnlaget fra Agresso og Sharepoint. Administrasjonen var positive til å bruke den ferdige BI-løsningen, da dette ville spare dem for mye tid. Vi opplevde de

som svært entusiastiske, og ivrige etter å hjelpe oss.

6.3.4.2 Hente inn data

Etter at dataen var på plass i databavarehuset, var det klart til at vi kunne bruke dataen i rapportene som Kundestyrte Prosjekt hadde laget. Vi måtte da først sette opp en kobling til databasen, før vi kunne skrive spørringer mot databasen for å hente ut data. Spørringene kunne gjøres i flere språk, blant annet SQL og MDX. SQL er et spørespråk for ER-databaser, mens MDX er et spørespråk for OLAP-databaser.

Spørringene som Kundestyrte Prosjekt hadde laget, ble skrevet i SQL, og passet kun til den databasen de hadde laget. Jeg og min kollega måtte derfor skrive nye spørringer, da både språket og databasen var endret. Nedenfor vises et eksempel på en spørring gjort i MDX. Spørringen henter inn data om registrerte timer, ansattes ansattenummer og navn, samt en oversikt over totalt antall fleksitid-timer de har tatt ut i 2009.

```
SELECT
    Measures.hours ON COLUMNS,
NON EMPTY
    Order(
        {[resource].[resource_nr].Members,
         [resource].[resource_name]. CurrentMember},
        Measures.[hours], ASC)
    ON ROWS
FROM
    Acandocube
WHERE
    ([activity].[activity_name].[Flexitid],
     [time].[year].[2009])
```

Figur 6.13: En av MDX-spørringene som ble laget i rapport-desingeren. Spørringen henter ut data fra kolonnen “hours” fra kuben “Acandocube”, hvor timene er fleksitid som er registrert i 2009.

Jeg og min kollega hadde kun vært borti SQL, og kunne dette godt. Vi brukte mye tid på å sette oss inn i MDX, og fikk se flere styrker med

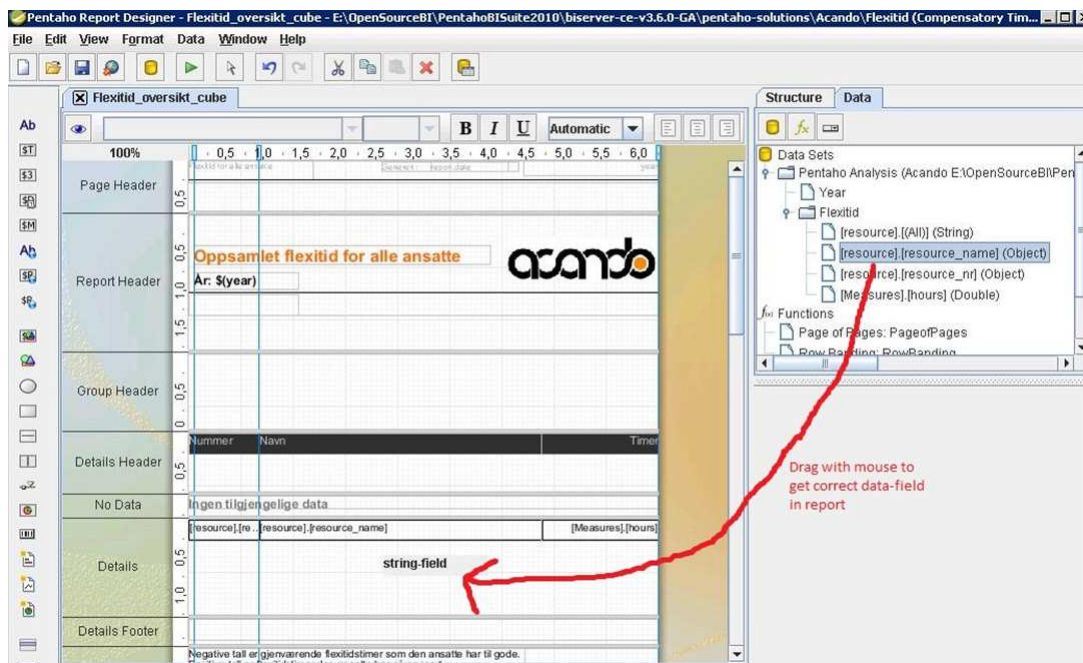
MDX-språket. MDX-spørringer var enklere enn SQL da det skulle hentes ut hierarki-spesifikk data. Dette kommer av at man i MDX spør direkte mot hierarki, noe som gjør det lettere å spesifisere hvilket hierarkisk nivå man vil hente data fra. Det var også svært enkelt å finne ønsket data når denne var lagret i et stjerneskjema, siden vi slapp å joine flere felt, slik som det gjøres i SQL. Utdrag fra min feltstudie-dagbok:

“SQL og MDX var tilsynelatende svært like språk i starten. Men der tok vi feil, for det er kun SELECT, FROM og WHERE som er likt i spørringene. Selve dataauthentingene er fundamentalt forskjellig, da SQL er designet for å kun spørre mot to dimensjoner (rad og kollonne), mens MDX ikke har noen grense for antall dimensjoner. Kollega X og jeg brukte flere dager på å prøve oss fram med disse spørringene.”

Feilmeldinger var svært dårlige i den delen av rapport-designeren hvor man skrev spørringer. Etersom MDX-editoren i kubedesigner ga gode feilmeldinger, valgte vi å skrive MDX-spørringer i kube-designeren, og heller kopiere disse over i rapport-designeren.

6.3.4.3 Rapport-design

Etter at vi hadde fått hentet inn data ved hjelp av spørringer mot databasen, var det klart for å designe rapportene. Den innhentede dataen ble da koblet opp mot feltene som skulle vises i rapporten, som vist i skjermbildet nedenfor.



Figur 6.14: Data som ble hentet inn gjennom spørringer mot databasen, ble koblet opp mot feltene i rapporten, som dynamiske parametre.

Kundestyrte prosjekter opplevde rapportdesigneren som et svært kraftig verktøy, som de måtte bruke en del tid på å lære seg. De beskrev verktøyet som et enkelt drag and drop-verktøy, med mange funksjoner, og de mente at verktøyet var lett å bruke når man først hadde lært seg det. Design av rapporter ble gjort ved å dra komponenter fra venstre-menyen over til rapport-ramma, og deretter sette tekst på disse. Teksten ble satt manuelt, eller ved å koble de opp mot parametrene fra innhentet database-data. De skrev i prosjektrapporten[25] at feilmeldinger var rapportdesignerens svake side, da disse var svært lite intuitive, og ofte var på formen “Det har oppstått et problem med komponent X” eller “Noe gikk galt når X skulle kjøres”. Dette medførte at de måtte bruke mye tid på å lete etter årsakene til feilene som oppstod.

Den mest komplekse rapporten som ble laget var “Personlig rapport”, som skulle vise ulike opplysninger om en gitt ansatt. Denne inneholdt data om gjenværende ferie, planlagt ferie med godkjent-status, uttatt flexitid, antall timer syk, tid brukt på egenopplæring og kurs, fakturerte timer, og gjennomsnittelig timepris fakturert til kunde. Ved å lage denne rapporten

ønsket Acando at ansatte skulle få muligheten til å se sin egen ansatt-profil, uten å måtte gå til administrasjonen og be om den. Kundestyrte Prosjekt opplevde dette som den mest komplekse rapporten av alle, ettersom den inneholdt mye forskjellig data, og flere komplekse utregninger og sammenligninger.

Gjenværende ferie

Angitt i dager. Rapport fra tidligere år kan være feilaktige i forhold til overført ferie

Ordinær ferie	Overført ferie	Sammenlagt ferie	Sammenlagt ferie etter planlagt ferie
25	0	25	25

Planlagt ferie

Fra dato	Til dato	Status	Antall arbeidsdager
14. sep, 2009	2. okt, 2009	Godkjent	15
1. nov, 2009	11. nov, 2009	Ikke behandlet	8

Flexitid

Angitt i timer

Gjeldende år	Sammenlagt
	-12
	-31,5

Sykdom

Angitt i timer

Egenmelding	Legeerklæring	Syke barn	Sammenlagt
0	0	0	0

Sykdom detaljer

Ingen registrerte data

Uke	Type	Antall timer
-----	------	--------------

Opplæring

Type	Antall timer
Egenopplæring	2
Kurs	16

Fakturering

Totale timer jobbet	Totale fakturerte timer	Faktureringsgrad %	Totalt fakturert
1 409,5	1 314	93,22	1 167 396,5

Timepris


Gjennomsnittlig timepris	Gjennomsnittlig inntjening per time
888,43	828,23

Figur 6.15: Samlet rapport for en enkelt ansatt. Ansatt-navn og -nr er her fjernet.

Kundestyrt Prosjekt-gruppa møtte store problemer da de publiserte rapportene fra Pentaho Report Designer til Pentaho-Serveren. De opplevde at rapportene ikke ble seende ut slik som de hadde gjort i designeren, og at noe data/informasjon manglet. Etter mye leting etter årsaken til dette, viste det seg at overlappende elementer i rapportdesigneren endte opp med å bli borte ved publisering til BI-Serveren. Dette skjedde uten at det kom noen feilmelding om problemet.

Kundestyrt Prosjekt utviklet etter hvert en template/mal for rapport-utseende, slik at de slapp å utvikle denne for hver nye rapport. De ble ikke klar over at man kunne lage slike maler, før et stykke ut i prosjektet. Et eksempel på en av rapportene med standard-utseende er vist nedenfor.

Planlagt ferie
År: 2009



Nummer: 00100
Navn: [Redacted]

Fra	Til	Status	Antall arbeidsdager
2009-09-14	2009-10-02	Godkjent	15
2009-11-01	2009-11-11	Ikke behandlet	8

Nummer: 00100
Navn: [Redacted]

Fra	Til	Status	Antall arbeidsdager
2009-09-18	2009-09-18	Ikke behandlet	1
2009-10-19	2009-10-19	Ikke behandlet	1
2009-11-12	2009-11-13	Ikke behandlet	2

Nummer: 00100
Navn: [Redacted]

Fra	Til	Status	Antall arbeidsdager
2009-11-09	2009-11-13	Ikke behandlet	5

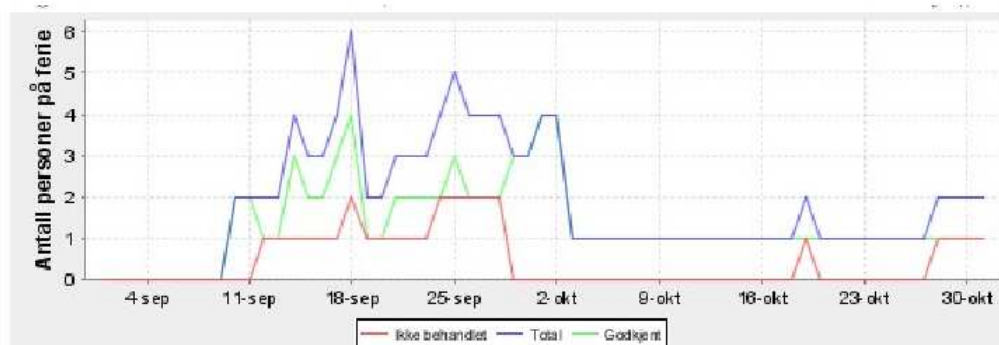
Nummer: 00100
Navn: [Redacted]

Fra	Til	Status	Antall arbeidsdager
2009-12-07	2009-12-11	Godkjent	5

Figur 6.16: Rapport som viser planlagt ferie for alle ansatte innen et gitt år. År velges i en drop-down boks øverst i rapporten.

Noen av rapportene fikk flere versjoner, f.eks. både en tekstlig rapport og en rapport med et kakediagram som illustrerte de samme tallene. De fleste rapportene som ble laget, hadde tidligere blitt laget manuelt av Administrasjonen, men også noen nye rapporter ble laget. Et eksempel på en av disse var en rapport som viste en samlet tidslinje over når de ansatte tok ferie, slik at man kunne se hvor mange som hadde ferie samtidig på et


gitt tidspunkt, som vist i rapporten nedenfor. Denne rapporten bruker den samme dataen, men fikk en annen type presentasjon.



Figur 6.17: Tidslinje som viser hvor mange ansatte som har tatt ut ferie til en hver tid i en gitt periode.

En annen ny rapport, viser en oversikt over hvor mye overtid ansatte kan jobbe, etter lovverket i arbeidsmiljøloven. Denne tillater administrasjonen å følge med på at de ansatte ikke jobber mer enn de har lov til. Denne rapporten ville vært svært tidskrevende å lage manuelt, og Acando så derfor stor verdi i den, da den ga verdifull informasjon, som de ikke hadde hatt tilgjengelig før.

Overtid oversikt
14. nov, 2009



NummerNavn	Overtid	Overtid	Overtid	Overtid	Overtid	Overtid	Maks	Maks	Maks	Maks	Maks
	side 2	side 3	side 27	side 42	side 47	side 50	overtid	overtid	overtid	overtid	overtid
	uker	uker	uker	uker	uker	uker	denne uken	ut	nesteneste 5	nesteneste 10	nesteneste 25
01 001	0	0	50	144	174	182	10	18	28	58	150
01 002	0	0	0	0	0	0	10	20	31	82	158
01 003	0	0	0	0	0	0	10	20	31	82	158
01 004	0	0	9	49	58	58	10	20	31	82	158
01 005	0	0	8	30	32	32	10	20	31	82	158
01 006	0	0	21	54	80	80	10	20	31	82	158
01 007	0	0	75	194	214	226	0	0	0	0	125
01 008	0	0	262	728	728	868	0	0	0	0	0
01 009	0	0	18	24	24	24	10	20	31	82	158
01 010	0	0	345	848	848	1 007	0	0	0	0	0
01 011	0	0	284	688	688	823	0	0	0	0	0
01 012	0	0	170	452	482	588	0	0	0	0	30
01 013	0	0	0	0	0	0	10	20	31	82	158
01 014	0	0	48	86	92	94	10	20	31	82	152

Figur 6.18: Tabell som viser hvor mye overtid en ansatt har lov til å jobbe i tiden framover.

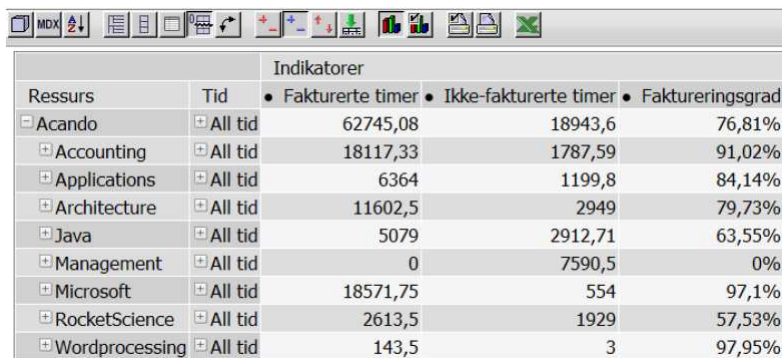
6.3.5 Analysering

Ved å ta i bruk OLAP-kube på toppen av databaseskjemaet, ble det mulig å analysere dataen i User Console gjennom web-browsersen. Vi lastet inn kubene til dette grensesnittet, og fikk opp alle dimensjonene på rekke, med målingen for faktatabellen i enden av Viewet. Dette ble mange dimensjoner, og mye informasjon på en gang. Vi skjønnte på dette tidspunkt at vi kunne ha laget flere kuber, i stede for å ha en for alle dimensjonene. Det var heldigvis mulig å velge bort dimensjoner fra Viewet, slik at det ble enklere å få oversikt.

Vi kunne gjøre drill-down-analyser på de ulike dimensjonene, ved å trykke på “+” og “-” for å navigere oss opp og ned i hierarkiene. På denne måten fikk vi se at målingene endret seg, slik at de passet det hierarkiske nivået. I figuren nedenfor vises fakturerte timer på de ulike prosjektene. Ved å trykke på et prosjekt, får man opp fakturerings-informasjon for hver enkelt ansatt på det prosjektet.

Jeg og min kollega opplevde analyse-verktøyet som intuitivt og enkelt å forstå. Utdrag fra dagbok:

“Vi fikk plutselig se flere sider av dataen som lå i databasen. Dersom jeg på tidsdimensjonen, og valgte 2009, januar, uke 3, og deretter valgte sykefravær fra aktivitets-dimensjonen, kunne jeg se antall sykefraværstimer for alle ansatte som var syke i uke



Ressurs	Tid	Indikatorer		
		• Fakturerte timer	• Ikke-fakturerte timer	• Faktureringsgrad
Acando	All tid	62745,08	18943,6	76,81%
Accounting	All tid	18117,33	1787,59	91,02%
Applications	All tid	6364	1199,8	84,14%
Architecture	All tid	11602,5	2949	79,73%
Java	All tid	5079	2912,71	63,55%
Management	All tid	0	7590,5	0%
Microsoft	All tid	18571,75	554	97,1%
RocketScience	All tid	2613,5	1929	57,53%
Wordprocessing	All tid	143,5	3	97,95%

Figur 6.19: Et skjermbilde av Analyse-grensesnittet i User-Consol. Dette illustrerer hvordan man kan navigere seg oppover og nedover i et hierarki.

3, 2009. Det begynner nå å gå opp for meg hvorfor det er blitt så populært med denne typen strukturering av data for analysering. Informasjon som krever avanserte spørringer mot databasen, blir på denne måten tilgjengelig for helt vanlige ansatte, med bare noen få tastetrykk”.

Andre analysefunksjonaliteter, som data mining og dashboards, ble ikke testet ut, da det var for lite tid til å sette seg inn i det. Vi noterte oss at Pentahos motor for mønster-gjenkjenning (Weka), fikk mye positiv omtale på ulike BI-forum. Dashboard-designeren finnes kun for Enterprise-versjonen, og vi fikk inntrykk av at å lage slike dashboards uten en editor ville bli svært vanskelig.

6.3.6 Aktørenes erfaringer og refleksjoner

Dersom man skal gjøre en vurdering av hvordan dette prosjektet har gått, er det nødvendig å ta i betraktning alle aktørenes erfaringer og meninger.

6.3.6.1 Kundestyrte Prosjekt

Prosjektet ble ikke helt som forventet, grunnet lite erfaring på BI-området. Studentene hadde sett for seg mer jobbing med rapport-design, og ikke så mye forarbeid som måtte gjøres først. Dette resulterte i en enorm feilestimering av den første sprinten, da de måtte gjennomføre hele ETL-steget før rapportutviklingen kunne starte.

Prosjektgruppa rapporterte i etterkant av prosjektet at de var fornøyde med jobben de hadde gjort, og at de trodde at Acando ville få god nytte

av rapportene de hadde laget, dersom de bestemte seg for å lage en full BI-implemterasjon [25]. Ettersom de ikke hadde tilgang på Acandos virkelig data fra AD, Agresso og Sharepoint, fikk de ikke bygget en database som passet denne dataen.

I retrospekt så prosjektgruppa noen ting de ville gjort annerledes, dersom de skulle gjort prosjektet om igjen [25]. De ville blant annet brukt en felles server for installasjon av Pentaho BI-Serveren og databasen. Denne hadde de installert på hver sin PC, og fikk derfor mye ekstraarbeid med å flette sammen all data, og koordinere arbeidet på slutten av hver sprint. De måtte ved flere anledninger endre på databasen, og dette krevde da at alle gjorde de eksakt samme endringene i sine lokale databaser. Mye tid kunne vært spart ved å bruke en felles server. De ville også ha utviklet en template/mal for hvordan Acando-rapportene skulle se ut, helt i starten av prosjektet, for å slippe å lage det samme designet flere ganger, og da risikere å at det ble ulikt.

Det ble gjentatte ganger i prosjektrapporten kommentert at det fantes lite dokumentasjon på de ulike verktøyene, og at feilmeldingene var gjen-nomgående dårlige. Dette mente de kunne være et resultat av at program-varen var Open Source-versjonen, og at dokumentasjon er noe man får dersom man kjøper Enterprise-versjonen av Pentaho. De mente likevel at den helt åpne versjonen av Pentaho var god nok for det arbeidet de hadde fått i oppgave å gjøre.

6.3.6.2 Sommeransatte

Jeg og min kollega hadde en fordel ved at vi hadde muligheten til å lese prosjektrapporten til Kundestyrt Prosjekt før vi startet sommerprosjektet. Vi kunne da unngå å gjøre de samme feilene, samtidig som vi fikk mange tips om hvordan installasjoner og konfigurasjoner burde gjøres. Vår veileder hadde også bedre kunnskap om BI, noe som hjalp oss mye under veis.

I starten brukte vi mye tid på å sette oss inn i BI som teknologi, og Pentaho som verktøy. Da vi leste Pentaho-dokumentasjonen var denne tilsynelatende hovedsakelig beregnet på utviklere med BI-kompetanse. Vi måtte derfor jevnlig søke på internett for å få BI-terminologier forklart.

Vi var svært fornøyde med de fleste programmene i Pentaho-plattformen, men vi fikk inntrykk av at brukervennligheten og feilmeldingene var av svært varierende kvalitet. Vi fikk ved flere anledninger erfare at deler av utviklingen ble vanskeligere med den helt åpne versjonen av Pentaho. Disse problemene var likevel ingen stor hindring, da det kun førte til at vi brukte lengre tid på enkelte oppgaver. Enterprise-versjonen ville vært nødvendig dersom

vi skulle laget dashboards. Ellers så det ut til å være et godt utviklingsmiljø på Pentaho-prosjektet, med mye aktivitet på forum og IRC-kanal, noe som var til stor hjelp når det oppstod problemer. Pentaho så ikke ut som et prosjekt som vil dø ut, da det har både mange ansatte og mange brukere.

I retrospekt av sommerjobben/feltstudiet følte jeg og min kollega at vi hadde fått en god innføring i BI som område, og Pentaho som verktøy. Denne kunnskapen var likevel kun overfladisk, og 6 uker var ikke lenge nok til å oppnå en trygghet på egen kompetanse og prestasjon. Vi var enige om at vi ikke hadde klart å gjøre denne jobben alene, selv om vi hadde fått dobbelt så lang tid. Vi så verdi i det å ha noen å diskutere med, spesielt når alt var nytt og fremmed, og vi skjønnte ting på ulikt vis. Vi så stor verdi i å ha veileder Y å støtte oss på, da han var til stor hjelp ilar utviklingen. Uten han ville vi kanskje ikke laget en ny database, og dermed endt opp med en lite optimal BI-implementasjon.

6.3.6.3 Oppdragsgiver/veileder

Veileder Y sier i etterkant av BI-prosjektet at Acando er svært fornøyde med det arbeidet som ble gjort, både av Kundestyrt Prosjekt og de sommeransatte. Prosjektet ble satt i gang av tre årsaker:

- For å implementere en løsning internt
- For å vurdere modenheten til Pentaho som BI-verktøy
- For å bidra med student-prosjekter, sommerjobber og Masteroppgaver

Ettersom implementasjonen ikke skulle gjøres av Acandos egne ansatte, men studenter, hadde ikke Acando veldig store forventninger til sluttresultat.

“Siden dette har vært et tredelt prosjekt, hvor læring har vært et viktig element, og studenter har vært involverte, har ikke forventningene vært så store. Basert på måten vi har kjørt dette prosjektet, har vi satt noen av utfordringene til side og fokusert på arkitektur og funksjonalitet i selve løsningen. Integrasjon mot live-systemene er ting som vi bevisst har satt på vent”.

Veileder Y mener resultatene har gått langt over forventningene, og at de er svært fornøyde med å ha satt i gang prosjektet i utgangspunktet. Det var Veileder Y som etterspurte en dimensjonal database-modell, og han sier seg i ettertid svært fornøyd med dette valget.

“Ved å holde fokus på arkitekturen og grunnmuren i løsningen, tilrettelegger vi for en god langtidslagring”.

Veileder Y forteller at Agresso er under oppgradering, og at det kanskje vil komme ny funksjonalitet der som vil være lik typisk BI-funksjonalitet. Det vil da gjøres en vurdering av Acando på hvorvidt Pentaho-implementasjonen skal satses videre på. Han er ikke fremmed for ideen om å investere i Pentaho Enterprise-versjonen i framtiden, dersom det skulle bli behov for det. Det er også ønskelig å benytte dashboards og data mining i en framtidig løsning. Veileder Y har helt fra starten av prosjektet, hatt et ønske om å få et dashboard med en lampe som lyser rødt, dersom noen ansatte jobber mer overtid enn de har lov til.

6.3.6.4 Administrasjonen

Administrasjonen i Acando har samarbeidet med begge utviklingsteamene som har jobbet med BI-implementasjonen. I starten kom de med ønsker om rapporter som kunne lages, og etter første rapportleveranse, ble rapportene evaluert av administrasjonen. I denne evalueringen kom det fram noen misforståelser som hadde oppstått under første iterasjon, og jeg og min kollega anså derfor denne evalueringen i fokusgruppen som svært verdifull for videre utvikling. De fleste tilbakemeldingene var utseende/design-relatert, men de hadde også noen kritiske tilbakemeldinger som gikk på tolkning av data-grunnlaget fra Agresso og Sharepoint.

Administrasjonen tror de vil spare mest tid på den personlige rapporten, da denne inneholder all informasjon som er av interesse for de ansatte. Denne vil utleveres automatisk til de ansatte på mail, enten ved et fast tidsintervall eller når ansatte forespør den.

“Det går mye tid til å grave i data, for å kunne svare på enkle spørsmål om timeføring og ferie. Med en slik personlig rapport slipper vi å bruke så mye tid på å lete etter denne informasjonen.”

Jeg og min kollega opplevde flere ganger at folk fra administrasjonen kom å spurte håpefullt om det var lenge til løsningen var ferdig. Det var åpenbart at de var positive til denne, og gledet seg over å kunne slippe den manuelle rapporteringen.

6.4 Vurdering

Pentaho BI Suite har vist seg å være en kraftig plattform, som gjør det mulig å bygge en BI-løsning som gir nye muligheter for bedriften. Plattformen

har alle nødvendige BI-verktøy for integrering, rapportering og analysing. Noen verktøy finnes kun i enterprise-versjonen. Acando har fått mulighet til å administrere ansatte og prosjekter på en helt ny måte. Manuelle prosesser automatiseres, og nye prosesser oppstår. Med disse vil det være lettere å administrere bedriften.

Verktøyene i Pentaho er av varierende kvalitet, og de ikke-Pentaho-spesifikke verktøyene (som Spoon) skilte seg positivt ut, med gode brukergrensesnitt og gode feilmeldinger. Rapportdesigneren utmerker seg også, med mange muligheter for å skreddersy rapporter, og med avanserte funksjoner for å gjøre analyser på innhentet data. Generelt framstod Pentaho som en svært kraftig plattform, med mange muligheter. Plattformens svakeste side var brukervennligheten, da den gjennomgående hadde dårlige feilmeldinger.

Den helt åpne versjonen av Pentaho bydde på tilsynelatende bevisste hindringer fra Pentahos side. Eksempler på dette er valget om å holde tilbake mye dokumentasjon for utviklere som bygger sine løsninger utelukkende på gratis-versjonen. På denne måten tjener Pentaho penger, ved at noen brukere ikke har nok med den lille dokumentasjonen som finnes for gratis-versjonen. Dersom Acando fortsetter sin Pentaho-implementasjon, vil de forsøke å implementere dashboards. Det vil da, høyst sannsynlig, bli nødvendig å oppgradere til Enterprise-versjonen, og Pentaho vil da tjene penger på Acando.

De største utfordringene som oppstod ved bruk av Pentaho-plattformen var relatert til mangel på kompetanse. Dette førte til feilestimeringer av tid, og at deler av implementasjonen ikke ble utført riktig og dermed måtte gjøres om igjen. Kundestyrte Prosjekt fikk et svært annerledes prosjekt enn de først trodde de skulle få, da de trodde at hovedjobben deres skulle bli å lage rapporter. Det har vist seg å være helt nødvendig å skjønne de grunnleggende prinsippene innenfor BI, før man starter en implementasjon. Da øker man kvaliteten på det man utvikler, og unngår å måtte rette opp feil senere.

Jeg og min kollega var fornøyde med å ha vært to på sommerjobben, da vi ikke så for oss dette som en enmanns-jobb. Dette fordi det var et helt nytt område, hvor det var nyttig å ha noen å diskutere med, for å forstå konseptene og prinsippene. Vi var også enige om at hele implementasjonen ville sett veldig annerledes ut dersom ikke Veileder Y hadde hatt erfaring med BI, og dermed vært i stand til å peke oss i rett retning.

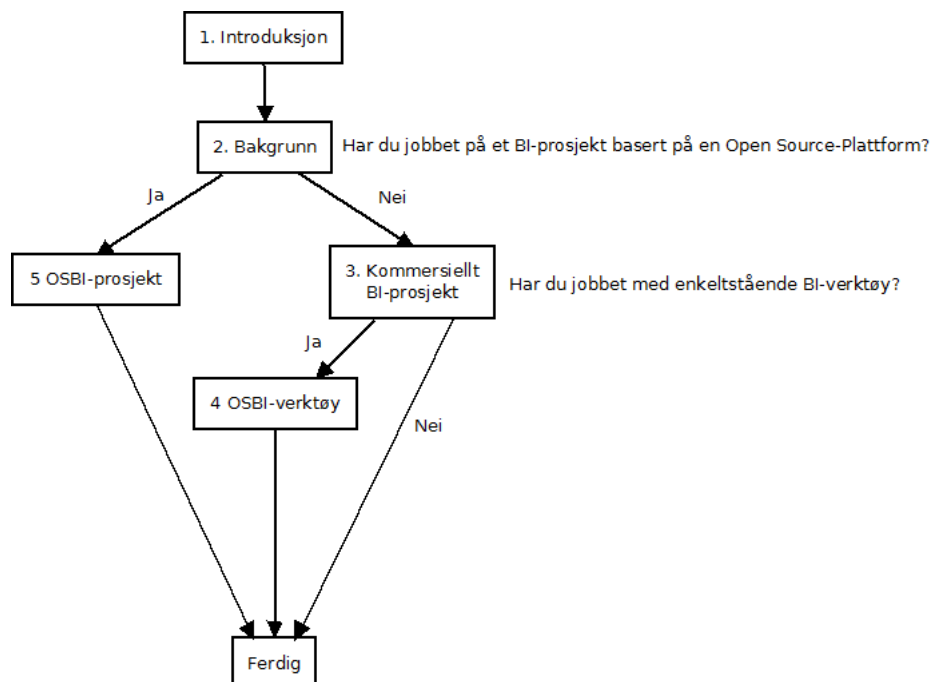
Kapittel 7

BI-bransjen

I etterkant av feltstudiet, ble det gjennomført en spørreundersøkelse, for å få et oversiktlig bilde over BI-bransjen. Det var ønskelig å se nærmere på SMB's behov for- og bruk av BI, og å sammenligne disse med tilsvarende for store bedrifter. Det var også ønskelig å undersøke motivasjoner for- og faktisk bruk av OSBI, i forhold til motivasjoner for- og bruksmønster av kommersiell BI. Dette ga et godt helhetsbilde over bransjen, og resultater som var ønskelig å se nærmere på. Jeg holdt derfor 3 intervjuer, med personer fra BI-bransjen. Dette ga dypere innsikt på noen av tema.

7.1 Spørreundersøkelse

Se appendiks A for en komplett liste med spørsmål fra spørreundersøkelsen. Undersøkelsen er delt inn i 5 deler, som angitt i flyt-diagrammet nedenfor. Diagrammet illustrerer hvordan respondentene ble ført gjennom undersøkelsen, basert på svarene som ble gitt. Alle fikk spørsmål om et BI-prosjekt de hadde jobbet på. Dersom respondenten hadde erfaring med OSBI, fikk han spørsmål om det siste OSBI-prosjektet han var på. Dersom han ikke har slik erfaring fikk han spørsmål om det siste BI-prosjektet han jobbet på, uavhengig av plattform. Det ble dermed samlet data om både OSBI-prosjekter og kommersielle BI-prosjekter, som da kunne sammenlignes i analysen. Det ble også spurt hvor mange brukere løsningen ville ha, for å få en viss ide om hvor stor bedriften som implementerte løsningen var. Dermed kunne jeg skille mellom svar fra SMB og store bedrifter.



Figur 7.1: Flytdiagram for spørreundersøkelsen.

Open Source-programmet Kwik Surveys [38] ble benyttet for utvikling av undersøkelsen. Ved å benytte Kwik Surveys hadde jeg alle nødvendige valg ved konstruering av spørsmål, og kunne velge mellom drop-down bokser, avkrysning av ett/flere alternativer, samt kommentarfelt der det var aktuelt.

Utviklingen av spørsmål til undersøkelsen ble dels inspirert av en BI-undersøkelse gjort av BARC¹[5], og dels basert på hypoteser som oppstod ila. feltstudiet.

18 norske konsultantselskaper med BI-kompetanse, fikk forespørsel om å delta på spørreundersøkelsen, og av disse sa 15 ja til å delta. Spørreundersøkelsen ble holdt aktiv i 21 dager², og på denne tiden var det totalt 74 personer som responderte.

¹Business Application Research Center

²Fra 24 februar til 10 mars, 2011.

7.2 Intervjuer

I etterkant av spørreundersøkelsen ble det holdt tre intervjuer. Målet med disse var å finne ut hvorfor overraskende funn var som de var, samt gå i dybden på interessante spørsmål. Intervjuobjektene holdes her anonyme, men beskrives i hver sin profil:

- Intervjuobjekt A: Seniordirektør i et av Norges ledende BI-selskap. Selskapet har om lag 900 ansatte totalt i Skandinavia, og intervjuobjektet har kun erfaring innenfor kommersiell BI.
- Intervjuobjekt B: Avdelingsleder for en BI-avdeling med 5 ansatte, i en bedrift på totalt 30 ansatte. Avdelingen benytter OSBI (Pentaho) til utvikling.
- Intervjuobjekt C: Daglig leder i en Norsk bedrift på 14 ansatte, som leverer datavarehus- og BI-løsninger. De har erfaring med OSBI, men er produktuavhengige i praksis.

7.3 Resultater

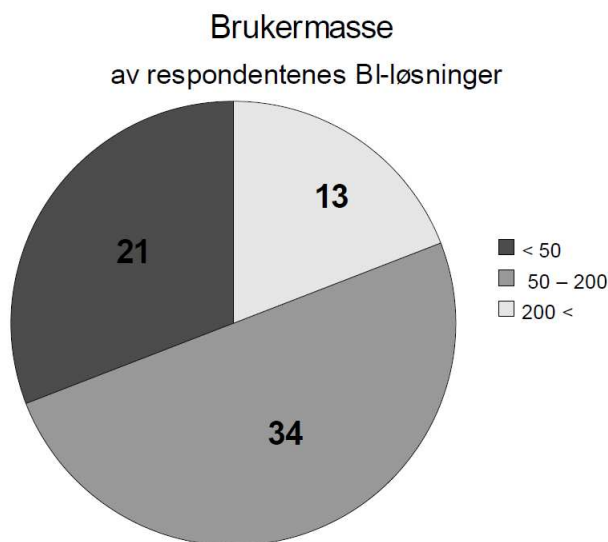
7.3.1 Respondentene

Det var totalt 74 respondenter i gjennomføringen av denne spørreundersøkelsen, og av disse hadde 10 jobbet med OSBI. En stor andel av respondentene hadde svært bred erfaring innenfor BI, både da det ble målt antall BI-prosjekter de hadde jobbet på og antall år de hadde jobbet med BI. 9 respondenter hadde jobbet på mer enn 15 BI-prosjekt, og av disse hadde 1 erfaring med OSBI.

7.3.2 SMB Vs. Store Bedrifter

BI-løsningene som respondentene fikk spørsmål om, hadde ulike antall brukere. Ved å se på antall brukere, får man et visst innblikk i hvor stor bedriften er, og inndelingen av SMB og store bedrifter er derfor basert på disse tallene³. I de fleste resultatene som sammenligner store bedrifter og SMB, er resultatene delt inn i kategoriene *små*, *mellomstore* og *store* bedrifter. Når SMB omtales, er det snakk om små- og mellomstore bedrifter slått sammen.

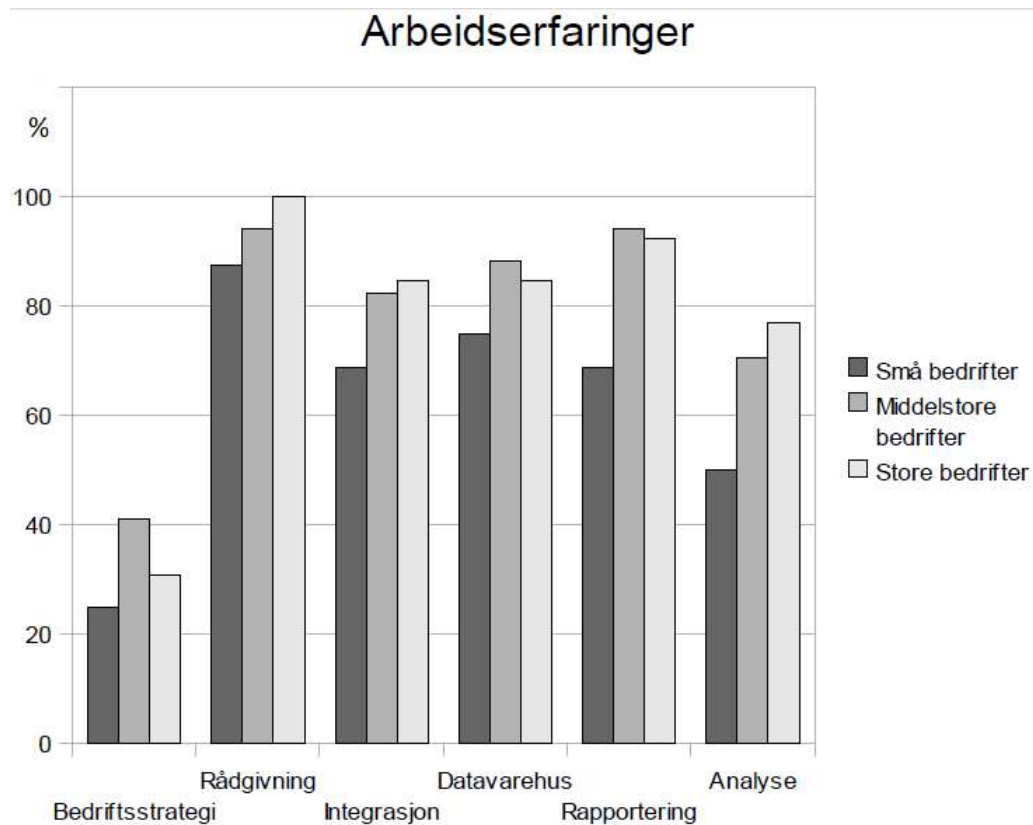
³Denne inndelingen kan være noe feilaktig, da store bedrifter også kan ha få brukere, avhengig av hva de ønsker å oppnå med BI.



Figur 7.2: Figuren viser hvor mange respondenter som jobbet med BI-løsninger for bedrifter med *små*, *middels store* og *store* mengder brukere.

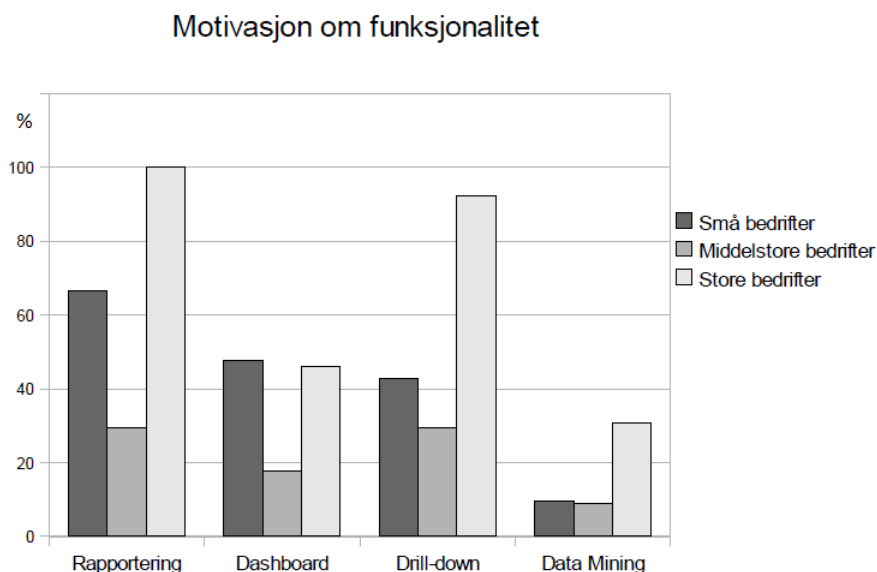
7.3.2.1 Erfaringsområder

Resultatene viser hvor mange prosent av de som jobbet på små, middels store, og store bedrifter, som hadde erfaring innenfor ulike BI-områder. Overraskende mange svarte at de jobbet med rådgivning, med tanke på at denne undersøkelsen ble uttrykt at var ment for tekniske utviklere og ikke rådgivere. Dette kan tyde på at mange enten har jobbet med begge deler, eller har jobbet med rådgivning tidligere. Det samme gjelder for bedriftsstrategi, da dette er mer forretningsrettet arbeid, og ikke direkte teknisk utviklingsarbeid.



Figur 7.3: Figuren viser hvor mange % av respondentene som hadde erfaring innen de ulike BI-områdene. Resultatene er sortert etter størrelsen på bedriften de jobbet for sist.

7.3.2.2 Motivasjon om funksjonalitet



Figur 7.4: Viser motivasjon om BI-funksjonalitet blant bedrifter av ulik størrelse.

Det kommer tydelig fram fra denne statistikken at motivasjon om bruk av BI-funksjonalitet preges av størrelsen på bedriften. Både store bedrifter og SMB fokuserer mest på rapportering og drill-down analyser, men de store bedriftene skårer likevel vesentlig mye høyere på begge deler. Store bedrifter ser ut til å generelt være mer motiverte av funksjonalitet, da de skårer høyest på samtlige funksjonaliteter. Mindre bedrifter motiveres kanskje av andre ting enn funksjonalitet.

SMB fokuserer vesentlig mindre på data mining enn de store bedriftene. Intervjuobjekt A mener dette kan komme av at mindre bedrifter ikke har den samme ROI av data mining, men at de heller får større utbytte av å bruke drill-down for å analysere datagrunnlaget. Han tror også at mindre bedrifter har større mangel på riktig kompetanse for å gjennomføre data mining.

“Det er lite kompetanse på dette området i Norge, og da spesielt hos SMB. Noen har kompetanse om det statistiske, men ikke selve bruken av verktøyene.”

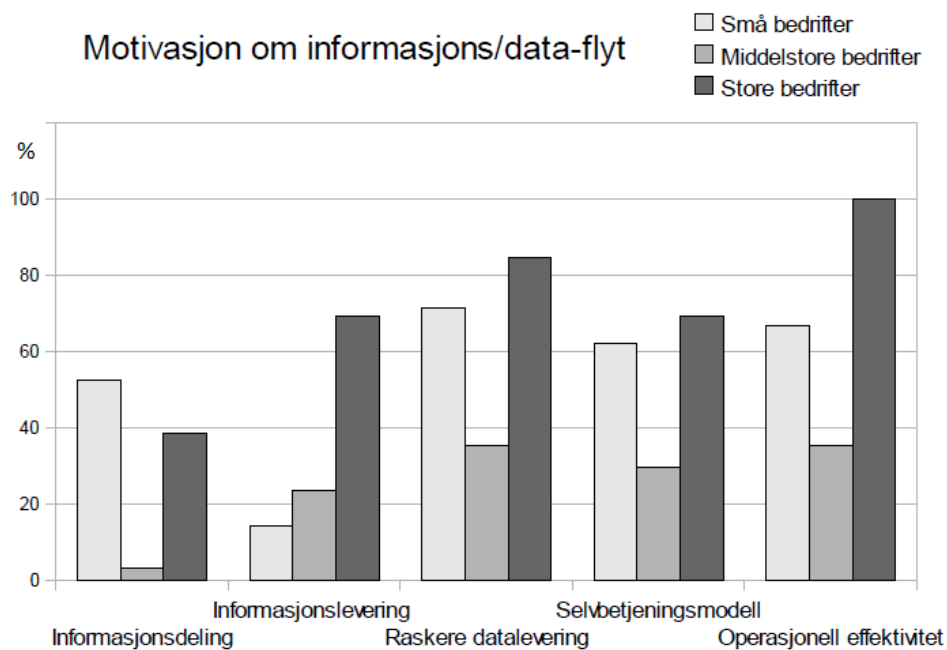
Intervjuobjekt B er enig med A i at det er mangel på kompetanse, og mener i tillegg at SMB ofte sliter med mangel på data, for dårlig datakvalitet, og dårlig struktur på den dataen som de faktisk har.

“I Norge er det ikke tillatt å lagre kundedata på personnummer. Dette er en ulempe når man skal sortere data om kunder fra flere systemer. Svenske kollegaer er kommet mye lengre på dette området, ettersom de lagrer kundedata på personnummer og dermed har en unik kunde-ID. De bruker da mye mindre penger på dataintegrasjon og får en bedre infrastruktur i løsningene sine.”

Intervjuobjekt B eksemplifiserer med et stort norsk prosjekt for en elektrokjede, som han deltok på. Prosjektet skulle bruke data mining for å analysere bruksmønstrene til kundene, men endte opp med å bli en fiasko, grunnet for dårlig infrastruktur og mangelfull kompetanse. B mener at data mining er mest brukt for CRM⁴/kunderelasjonshåndtering, og at de fleste SMB-ene ønsker å bruke BI til andre ting enn kundefølgning. Intervjuobjekt A mener at selv de som ønsker å benytte data mining, står overfor mange hindringer/utfordringer.

⁴Customer Relations Management

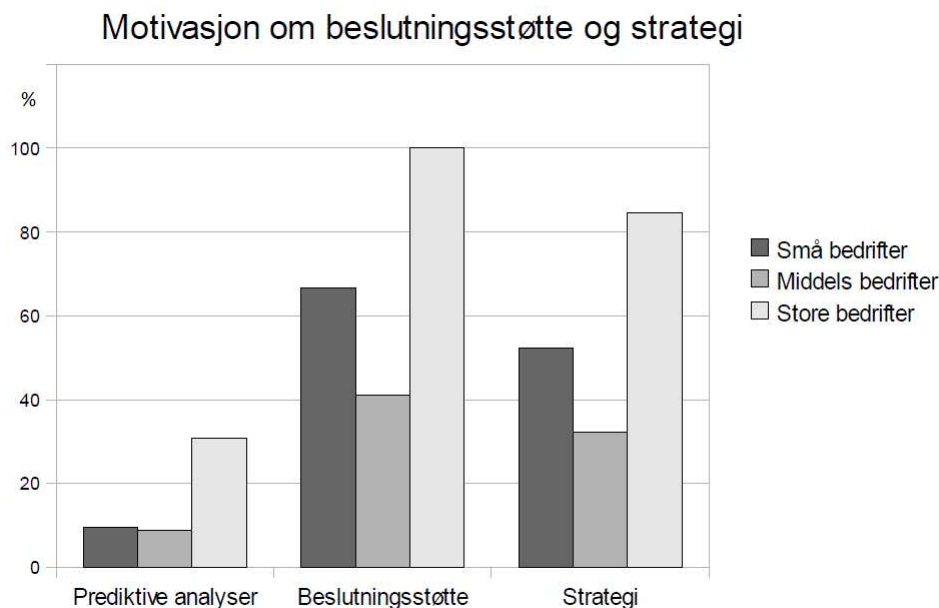
7.3.2.3 Motivasjon om informasjons-/dataflyt



Figur 7.5: Statistikken viser motivasjoner om informasjons/data-flyt blant bedrifter med små, middels store og store brukergrupper.

Her skårer de store bedriftene høyere på samtlige punkter. SMB fokuserer mest på rask datalevering, operasjonell effektivitet og å oppnå en selvbetjeningsmodell. Samtlige av disse tre punktene er relatert til rapportering, og som det kom fram i foregående statistikk, så er det nettopp rapportering som er viktigst av funksjonalitet for SMB.

7.3.2.4 Motivasjon om beslutningsstøtte og strategi



Figur 7.6: Viser motivasjoner om beslutningsstøtte og strategi blant bedrifter med små, middels store og store brukergrupper.

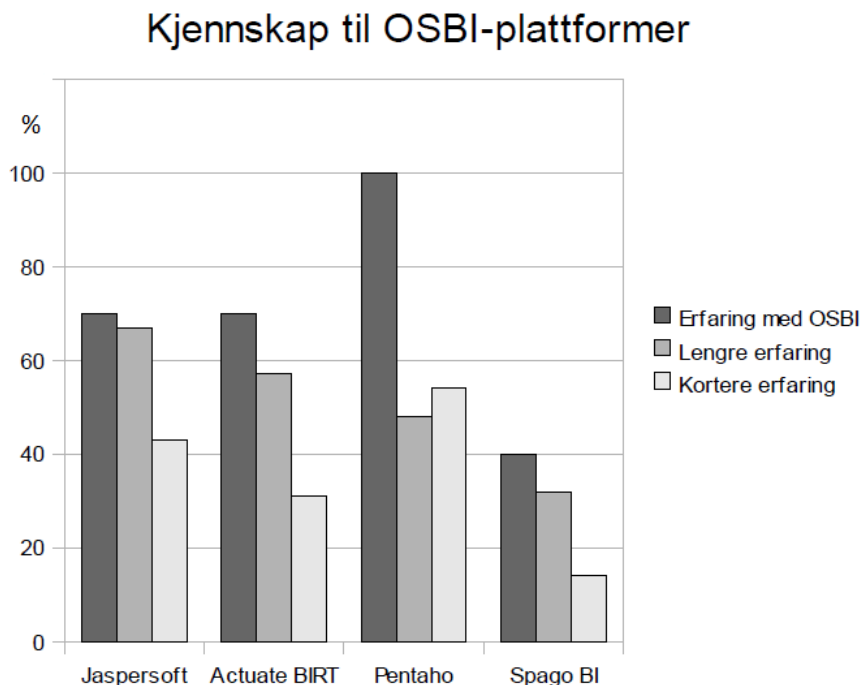
Også på beslutningsstøtte og strategi ser det ut til at de store bedriftene har et høyere fokus. Intervjuobjekt C tror at en naturlig innfallsvinkel for SMB-er til BI, er gjennom rapportering; de har typisk en rapport som de ønsker å få mer ut av. Han mener at SMB helt klart har andre behov enn de store bedriftene.

“Behovet for BI vokser med bedriften, og de fleste SMB-er er ikke modne nok for strategi og forretningsstøtte. “

7.3.3 OSBI vs. kommersiell BI

10 av de 74 respondentene hadde erfaring med OSBI. Dette tallet er preget av at jeg aktivt forsøkte å finne respondenter med OSBI-erfaring, og ville høyst sannsynlig vært vesentlig lavere dersom jeg ikke hadde tatt hensyn til respondentenes bakgrunn i oppsøkningsprosessen.

7.3.3.1 Kjennskap til de største OSBI-plattformene



Figur 7.7: Figuren viser hvor mange som kjente til de ulike OSBI-plattformene. Svarene er delt inn i de som har jobbet med OSBI, de som har jobbet med BI i mer enn 15 år, og de som har jobbet med BI i mindre enn 15 år.

Figuren over viser kjennskap til OSBI-plattformene, som er målt for tre ulike grupper; de som selv har erfaring med en eller flere OSBI-plattformer, de som har jobbet med BI i mer enn 15 år, og de som har jobbet med BI i mindre enn 15 år. Her skårer de to første gruppene overraskende lavt. Jeg hadde forventet at de som var i OSBI-miljøet hadde kjennskap til hvilke andre OSBI-plattformer som eksisterte, men det var kun Pentaho av de 4 plattformene, som var kjent av alle med OSBI-erfaring. Dette kan komme av at 9 av de 10 respondentene selv hadde jobbet med Pentaho. Det var også forventet bedre kjennskap til OSBI-plattformene blant de som hadde jobbet i BI-bransjen i mer enn 15 år, men statistikken viser at kjennskap til de 4 plattformene lå på mellom 35 % og 70 %. Jeg forventet også at de som hadde lengre erfaring i bransjen skulle ha bedre kjennskap til plattformene enn de

som hadde kortere erfaring, og dette var tilfelle for 3 av de 4 plattformene.

Intervjuobjekt A mener at dette dårlige kjennskapet, kommer av at OSBI-leverandører er svært dårlige til å markedsføre seg. Han bruker en “isfjellsanalogi” for å forklare dette.

“Toppen av isfjellet er det du ser. Det norske markedet er svært funksjonsorientert, og liker at det ser pent ut. Mange kommersielle leverandører er flinke til å markedsføre seg på denne fronten, og blir da toppen av isfjellet; det vi ser. OS er ofte dårlig til å markedsføre seg, og faller derfor typisk nedenfor det du ser av isfjellet.”

A tror videre at det er litt enten eller med OSBI; enten er du i det rette miljøet hvor dette er interessant, eller så faller OSBI helt utenfor interesseområdet. Han mener at det finnes så mange BI-plattformer på markedet at det er de man hører ofte om som blir lagt merke til og husket.

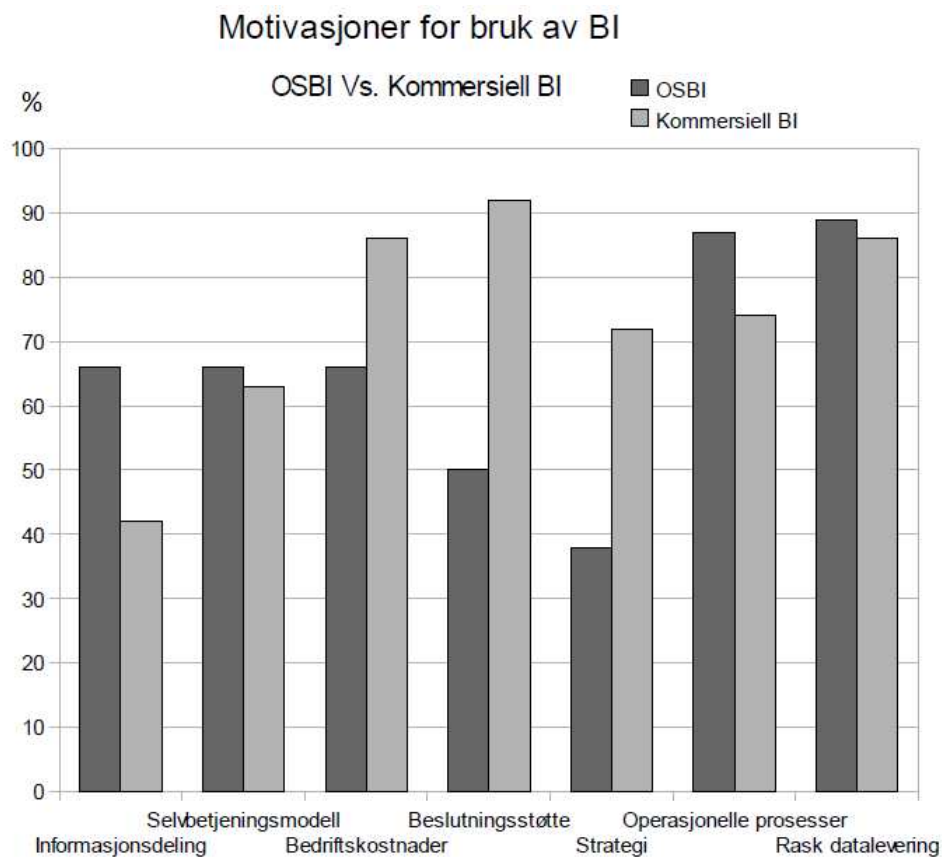
7.3.3.2 Antall brukere

Av de 10 som hadde jobbet på OSBI-prosjekter rapporterte alle mindre enn 100 brukere for BI-implementasjonene. Gjennomsnittelig rapporterte de 40 brukere, sammenlignet med 366 brukere for prosjekter med kommersielle BI-plattformer⁵. Dette betyr at samtlige bedrifter som benyttet seg av OSBI, var SMB-er.

7.3.3.3 Motivasjoner for bruk av BI

Ved å se på motivasjoner for investering av BI, får man et bilde over selve årsaken(e) til at man tok steget mot å investere i en slik løsning. Nedenfor vises hva respondentene svarte, da de ble spurt om hva kundens motivasjoner for investering i BI var. Resultatene er sortert på respondenter med OSBI-erfaring, mot de som ikke hadde slik erfaring.

⁵Tallene er regnet ut ved å ta medianen av spekteret brukere i hver valgte gruppe. Medianen ble ikke regnet ut på alternativet som het 5000+ brukere, da dette var umulig å regne ut. Kun en person valgte dette alternativet, hvilket gjør det rimelig å anta at reliabiliteten ikke ble svekket betydelig.



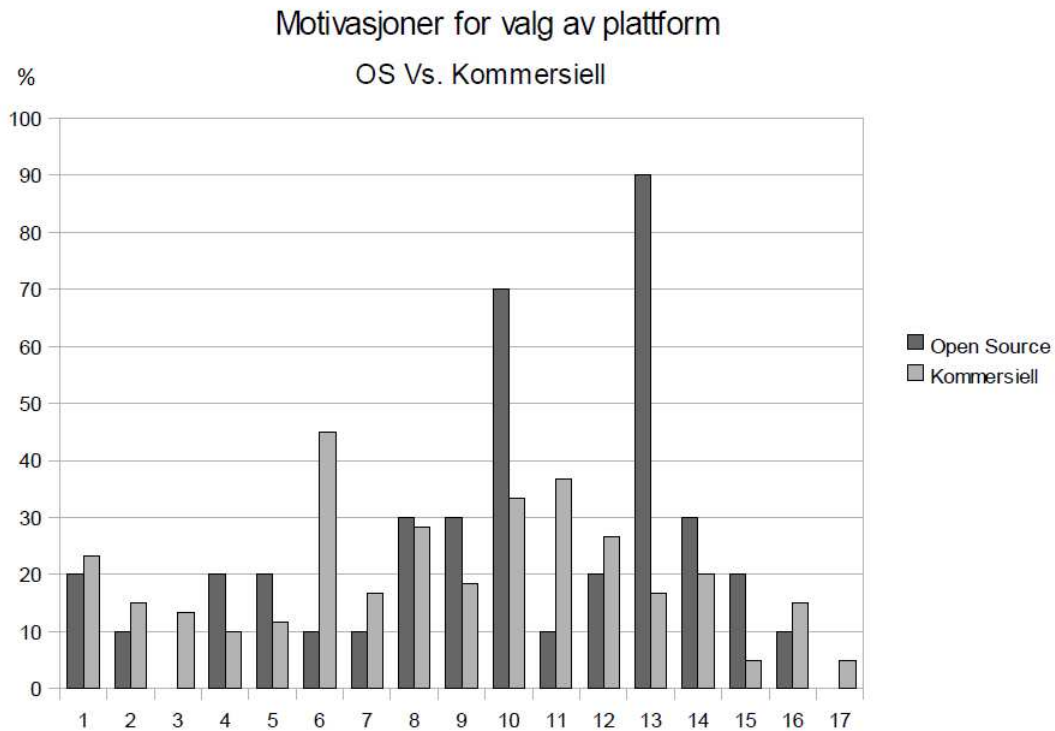
Figur 7.8: Motivasjonsfaktorer for investering i BI, sortert på bedrifter som brukte OSBI vs. kommersiell BI.

I diagrammet ovenfor er det tre av punktene som utmerker seg som varierende mellom bedrifter som bruker OSBI og bedrifter som bruker kommersiell BI. Første punkt, informasjonsdeling, var viktig for 72 % av OSBI-bedriftene, men bare 38 % av de kommersielle. På punkt fire kommer det fram at de kommersielle var langt mer motiverte av beslutningsstøtte enn de som benyttet OSBI (91 % vs. 50 %). Det siste nevneverdige punktet, punkt 5, er motivasjonen om bedre bedriftsstrategi, hvor 72 % av de kommersielle svarte at dette var viktig, mens bare 29 % av OSBI mente at dette var viktig. De to siste punktene, hvor kommersiell BI skårte langt høyere enn OSBI, har til felles at det handler om bruk av BI for forretningsstøtte (strategi og forretningsbeslutninger).

7.3.3.4 Valg av BI-plattform

På ett av spørsmålene fikk respondentene spørsmål om å velge opptil tre årsaker til valg av BI-plattform. De kunne da velge mellom følgende:

1. Evnen til å støtte et stort antall brukere samtidig
2. Tilgjengeligheten for lokal støtte
3. Hørte til et annet produkt
4. Leverandør gjorde en god salgsjobb
5. Fullførte ”proof of concept” raskere, eller bedre enn andre
6. Bedriftsstandard
7. Brukervennlighet for utviklere
8. Brukervennlighet for sluttbrukere
9. God ytelse
10. Funksjonalitet/produktfunksjoner
11. Integrerer med andre produkter allerede i bruk
12. Evne til å håndtere store datamengder
13. Lav pris
14. Produktrykte
15. Utvalg av støttede server-plattformer
16. Forhold til leverandør
17. Web-arkitektur



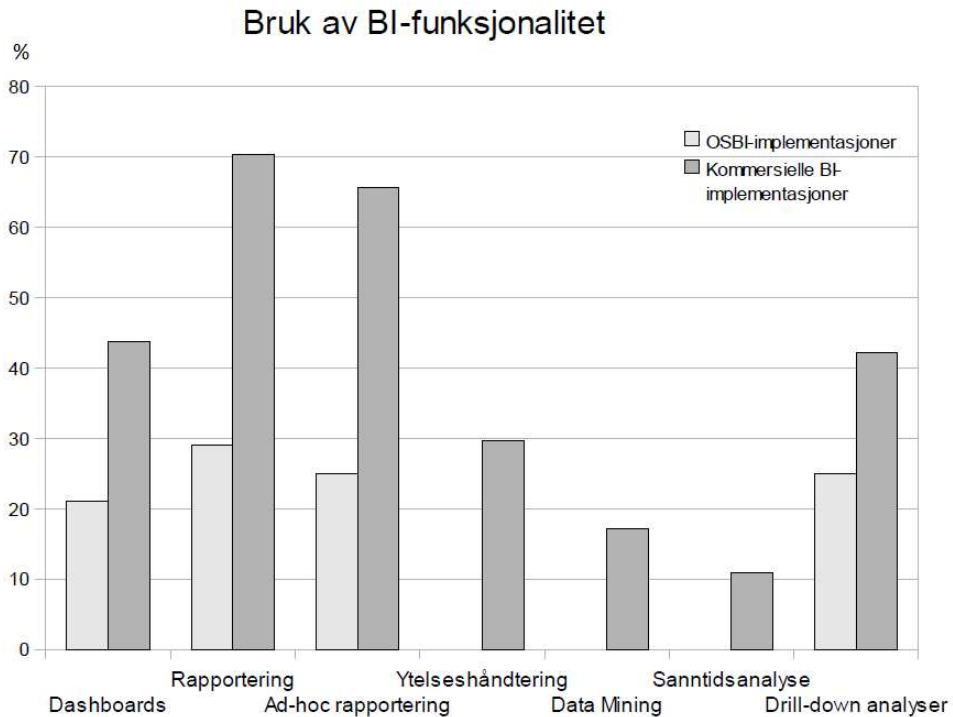
Figur 7.9: Motivasjoner for valg av BI-verktøy. Resultatene viser hvor mange % som svarte hva. Hver respondent kunne velge inntil tre alternativer.

OSBI ble oftest valgt grunnet lav pris og funksjonalitet, mens kommersiell BI oftest ble valgt fordi det er en bedriftsstandard og var integrert med andre produkter. De to årsakene til valg som skårte høyest for kommersielle plattformer er ikke direkte *valg*. Når man bruker en bedriftsstandard er dette ikke et valg man tar; det er slik det allerede gjøres i den bedriften. Dersom plattformen er integrert i andre produkter, er heller ikke dette et direkte valg, da plattformen allerede var bestemt før de bestemte seg for å bedrive BI. Etter disse to “ikke-valgene” var funksjonalitet, og store datamengder, de vanligste motivasjonsfaktorene for valg av kommersielle plattformer.

Valg av OSBI-plattform var i 90 % av tilfellene motivert av lav pris, og 70 % av tilfellene motivert av tilgjengelig funksjonalitet. Valg av OSBI og kommersielle plattformer har til felles at de i stor grad motiveres av funksjonalitet, men de som velger kommersiell BI gjør det likevel dobbelt så ofte pga. funksjonalitet. Dette kan tyde på at de som valgte OSBI generelt var mindre fokuserte på funksjonalitet enn de som valgte kommersiell BI.

OSBI ble i mye større grad valgt grunnet lav pris. Dette kan være et resultat av at de som velger OSBI-plattformer har lavere omsetning enn de som velger kommersiell BI. Det er overraskende at det basert på rykte er flere som velger OSBI enn kommersiell BI (40 % Vs. 30 %). Valg av kommersiell BI og OSBI er like motivert av brukervennlighet for sluttbrukere, men ikke for brukervennlighet for utviklere, da kommersiell BI oftere velges på grunn av dette.

7.3.3.5 Bruk av BI-funksjonalitet



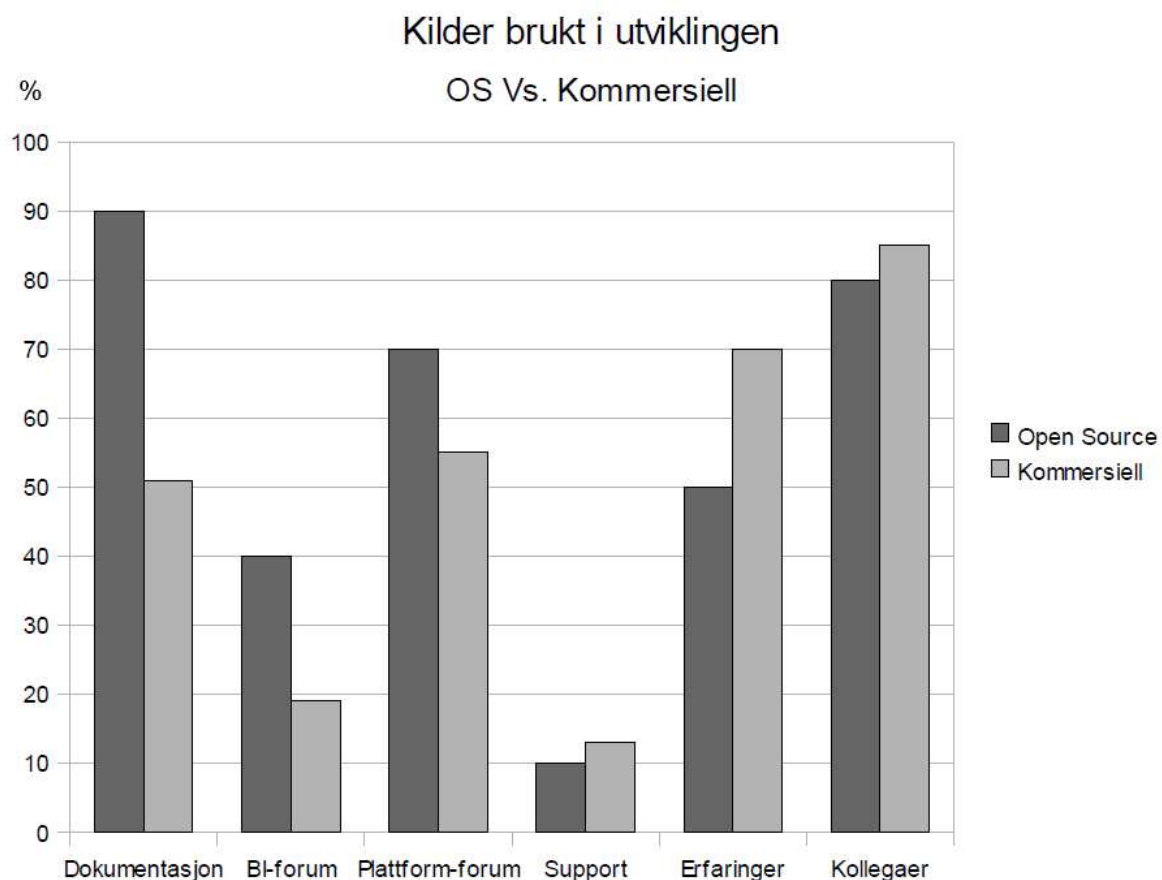
Figur 7.10: BI-funksjonalitet benyttet, sortert på bedrifter som benyttet seg av OSBI Vs. kommersiell BI.

De bedriftene som gjorde BI-implementasjoner basert på OSBI tok i bruk langt mindre funksjonalitet, sammenlignet med de som benyttet kommersiell BI. Mellom 20 og 30 % av OSBI-implementasjonene tok i bruk dashboards, rapportering, ad-hoc rapportering og drill-down analyse. Variasjonen på antall prosent er svært liten i det de bruker, mens tre av funksjonsområdene

ikke er brukt i det hele tatt; ytelseshåndtering, data mining og sanntidsanalyse. For de kommersielle BI-implementasjonene var spredningen større i brukt funksjonalitet. Høyest skårte vanlig- og ad hoc-rapportering, etterfulgt av dashboards og drill-down analyser. De mest brukte var altså de samme som for OSBI, men i tillegg brukte om lag 30 % av de kommersielle BI-implementasjonene også ytelseshåndtering. 18 % brukte data mining, mens 11 % brukte sanntidsanalyse.

7.3.3.6 Kilder brukt under utviklingen

Ved å se på støttekilder brukt under utviklingen, kan man få et innblikk i hvordan utviklerne går fram for å løse problemer/utfordringer under implementasjonen.



Figur 7.11: Oversikt over hvor mange prosent av utviklerne som svarte at de brukte gitte kilder ofte/svært ofte i utviklingen.

Det ser ut til at utviklere som jobber med OSBI i større grad benytter seg av begge typer forum. Det var overraskende at OSBI-utviklere også i større grad enn kommersiell BI-utviklere benytter seg av dokumentasjon til verktøyet brukt. Resultatene var forventet motsatt, da jeg opplevde dokumentasjonen til OSBI-verktøyet Pentaho som svært dårlig. Siden 5 av 10 OSBI-respondenter brukte Enterprise-versjonen, og da hadde tilgang til mer dokumentasjon, kan dette være årsaken til at disse utviklerne brukte dokumentasjon i så stor grad.

Utviklere som jobbet med kommersiell BI støttet seg mest på kollegaer og erfaringer, mens OSBI-utviklere støttet seg mest på dokumentasjon,

kollegaer og plattform-forum. OSBI-utviklere brukte kundestøtte/support i nesten like stor grad som utviklere på kommersiell BI, noe som var overraskende, siden support kun følger med enterprise-versjonen av OSBI-plattformen. Intervjuobjekt C mener at support i enkelte tilfeller er ekstremt viktig, og at man er nødt til å bruke Enterprise-versjonen av OSBI dersom man skal lage en seriøs implementasjon basert på OSBI. Han var skeptisk til at det kunne finnes noen i Norge som hadde laget en seriøs OSBI-implementasjon, uten å bruke enterprise-versjonen, grunnet behovet om support.

7.3.3.7 utfordringer for OSBI

Intervjuobjekt B mener at OSBI er lite kjent fordi det er nytt på markedet, og at det er lite brukt fordi det kreves en spesiell type kompetanse for OSBI-utvikling.

“Man har mindre definerte utviklingsroller, og må hele tiden fikse ting sjøl; ting som man ikke har vært borti før.”

Han mener videre at man ved bruk av OSBI får en billigere pris i produktanskaffelse, men at man likevel får store utgifter i konsulenttimer, eller timer for egne IT-ansvarlige, for å gjennomføre implementasjonen. Man må derfor ha et betydelig budsjett uansett, og får sannsynligvis ikke en billig BI-implementasjon, selv om programvaren er gratis. Intervjuobjekt C peker på en annen utfordring relatert til kompetanse.

“Dersom du setter opp en OSBI-implementasjon og mister de ansatte som gjorde implementasjonen, vil du kanskje få problemet med å finne personer som har den riktige kompetansen. Det er få personer med OSBI-kompetanse, både i verden og i Norge. Dersom du derimot jobber med en av de store leverandørene, som SAP, kan du enkelt leie utviklere fra Capgemini eller Accenture til å jobbe med BI-løsningen din.”

C mener videre at en forutsetning for å kunne lykkes med OSBI er å ha en intern entusiasme for programvaren, og ha en kultur for bruk av Open Source. I bedriften hvor C jobber, forsøker de å “avmystifisere” OS, og fronter OS på lik linje med kommersiell programvare. C mener at BI-programvare alene ikke er i stand til å ødelegge et BI-prosjekt.

“BI handler jo egentlig lite om software. Det handler mest om arbeidsprosesser, forretningsregler, virksomhetsstrategier og beslutningsmønstre som skal underbygge strategier. Det er derfor sjelden

software ødelegger et prosjekt. Det klarer som regel konsulenter eller andre fint på egen hånd. Undervurdering av kompleksitet i KPI-er, dårlig datakvalitet, mangelfull kompetanse om arkitektur og modellering, samt for liten involvering av forretnings siden er klassiske fallgruver for BI prosjekter.”

I selskapet hvor A jobber, er holdningen til Open Source noe annerledes. Selskapet mener at investering i programvare er et viktig steg mot en god BI-implementasjon.

“Det er ikke ønskelig å ta sjansen på å bruke en Open Source-programvare som er basert på et utviklingsprosjekt som kan komme til å dø ut. Vi ønsker ikke å ta ekstra risiko ved valg av programvare, men velger heller noe sikkert innenfor teknologien. På denne måten viser vi kundene våre at vi tar “sjenerøse” valg, og tar utviklingen seriøst.”

7.3.3.8 OSBI i framtiden

9 av de 10 respondentene som hadde OSBI-erfaring, mente at OSBI-bedriftsmodellen er levedyktig på lang sikt. Intervjuobjekt A tror OSBI kan bli mer brukt i framtiden, dersom kommersiell BI ikke går ytterligere ned i pris.

“Kommersiell BI koster i dag halvparten av hva det gjorde for fem år siden. Programvaren blir bare billigere og billigere, og du får i dag mye BI for pengene. Jeg ser for meg at OSBI kan presse prisene ned ytterligere. Men dersom prisene ikke går mer ned kan kanskje OSBI vinne terreng i BI-markedet, slik som for eksempel Open Source-operativsystemet Linux har gjort innenfor operativsystem-markedet.”

7.3.3.9 OSBI for SMB

10 av 10 respondenter med OSBI-erfaring, mente at OSBI kan være et godt alternativ for SMB-er, mens 6 av 10 mente at OSBI kunne være et godt alternativ for store bedrifter. Intervjuobjekt C mener at OSBI er godt nok for enkelte SMB-er.

“Vår tilnærming til OSBI er at vi anser dette som “modent nok” til å tas i bruk av bevisste virksomheter som innehar høy intern kompetanse, og som i tillegg har et oversiktlig miljø. Med dette mener vi at OSBI ikke kan erstatte tradisjonelle leverandører hos

store virksomheter som Telenor, DNB Nor, SPK, NAV osv, men hos mindre virksomheter som har færre kilde-systemer og mer homogene brukergrupperinger.”

C mener altså at bedrifter som gjør OSBI-implementasjoner er avhengige av å ha intern kompetanse i bedriften. A er enig i at intern kompetanse er viktig, men tror at SMB har for lite personell til å holde denne kompetansen. A stiller seg skeptisk til bruk av OSBI, også for SMB.

“Hvis programvare er problemet, er du feil ute. Hvis du først har bestemt deg for at du trenger BI, burde du kjøre trygt.”

7.4 Funn

7.4.1 SMB's bruk av BI

I følge resultatene om motivasjoner for BI-implementasjoner, kom det fram at SMB-er generelt har lavere fokus på funksjonalitet, sammenlignet med de store bedriftene. SMB-er motiveres mer av raskere datalevering, operasjonell effektivitet og en selvbetjeningsmodell for brukerne. Av det funksjonelle har SMB mest fokus på rapportering og drill-down analyser, og minst fokus på data mining, strategi og beslutningsstøtte. I følge flere av intervjuobjektene kan dette forklares med at SMB-er har mindre datagrunnlag, dårligere datakvalitet, og mangel på kompetanse. Mindre bruk av data mining kan ha flere årsaker:

- Et generelt mindre behov, da SMB i mindre grad håndterer CRM-systemer
- For lite datagrunnlag til å kunne lære noe av det
- Mangel på standardisert kunde-ID
- For dårlig datakvalitet
- For dyrt
- Mangel på kompetanse
- Drill-down analyser gir større ROI enn data mining

7.4.2 OSBI

Kjennskap til OSBI-plattformer var overraskende lav. De som hadde jobbet med OSBI hadde best kjennskap til de ulike OSBI-plattformene, etterfulgt av de som hadde jobbet med BI i mer enn 15 år. Den dårlige kjennskapen kan skyldes flere ting; OSBI-leverandører er dårlige til å markedsføre seg selv, OSBI er nytt på markedet, og man må være i OSBI-miljøet for å ha en interesse for disse plattformene.

Bedrifter som velger OSBI er ofte motiverte av den lave prisen på programvaren, mens bedrifter som velger kommersiell BI i stor grad motiveres av å være i stand til å håndtere store datamengder.

Kompetanse om OSBI er ikke lett å få tak i blant norske konsulentfirmaer. Dette kan øke terskelen for å bruke OSBI, da det kan bli vanskelig å få hjelp til utbedringer av en OSBI-implementasjon. Selv om man bruker OSBI betyr ikke dette nødvendigvis at implementasjonen blir billigere enn ved bruk av kommersiell BI, da det uansett er konsulenttimene som vil koste mest. Ved bruk av et vanskelig verktøy vil det bli brukt flere konsulenttimer til implementasjonen, sammenlignet med bruk av et enkelt/brukervennlig verktøy. Resultatene viste at de som valgte OSBI ikke hadde et like stort fokus på brukervennlige utviklingsverktøy, som de som valgte kommersielle BI-verktøy.

7.4.3 OSBI for SMB

Ved å sammenligne motivasjoner for bruk av OSBI med fokusområdene til SMB-ene som deltok i undersøkelsen, ser det ut til å være et stort overlapp. Dersom SMB-enes fokusområder samsvarer med SMB-enes behov, er det rimelig å anta at OSBI i stor grad vil oppfylle SMB-enes behov. Rapportering vha. en selvbetjeningsmodell et vanlig fokusområde for SMB, som helt klart er mulig å implementere ved hjelp av OSBI.

I følge intervjuobjektene burde SMB-er som gjør BI-implementasjoner basert på OSBI, ha gitte forutsetninger. Bedriften burde ha få kildesystemer som skal integreres, og homogene brukergrupperinger. Bedriften må ha høy intern kompetanse, da det ikke er lett å få ekstern hjelp. Denne kompetansen må være på flere plan; BI-arkitektur, databasemodellering, estimering, integrasjon, og BI-funksjonalitet. Bedriften må ha en god kommunikasjon mellom de som tar forretningsavgjørelser og de som utvikler det tekniske. Det burde også eksistere en intern entusiasme for bruk av OSBI, samt en generell kultur for bruk av Open Source. Flere av intervjuobjektene mente at man burde benytte Enterprise-versjonen, for å få support og doku-

mentasjon tilgjengelig. Uten denne vil det kanskje bli brukt mye lengre tid på implementasjonen, og det kan bli vanskeligere å få hjelp raskt, i kritiske situasjoner.

7.4.4 OSBI i framtiden

Prisene på kommersiell BI er halvert ilt. de siste 5 årene, og markedet forventer at denne prisnedgangen vil fortsette. OSBI vil kanskje vinne terreng, og bli brukt i større grad i framtiden, dersom prisene på kommersiell BI ikke går ytterligere ned.

Kapittel 8

Diskusjon

Dette studiet ble gjennomført for å undersøke hvorvidt en OSBI-plattform vil være tilfredsstillende for SMB-er som ønsker å benytte BI. Det ble ila. studiet tydelig at det ikke bare er programvare som er avgjørende for at SMB skal lykkes med BI, men at det også finnes andre utfordringer som det må overkommes. Problemstillingen ble derfor utvidet til også å omhandle generelle utfordringer for SMB-er som ønsker å implementere BI, samt hvordan disse burde håndteres.

Det er tidligere gjort lite forskning med konkrete eksempler på BI-implementasjoner hos SMB-er, og da spesielt ved bruk av OSBI. Den forskningen og teorien som er skrevet på dette området, er stort sett på bransje-nivå, noe som fører til at dette studiet dessverre ikke har noe direkte sammenligningsgrunnlag når feltstudiet gjort på Acando skal diskuteres.

Diskusjonen er delt inn i tre seksjoner; første seksjon diskuterer SMB's behov for BI, og utfordringer relatert til selve bruken. Seksjon to diskuterer OSBI's modenhet. Siste seksjon forsøker å foreslå hvordan SMB-er burde gå fram når de skal implementere BI, samt hvilke typer bedrifter som kan vurdere å benytte OSBI.

8.1 SMB's behov for BI

8.1.1 Automatisering av manuelle prosesser

Gartner [27] og Wise [76] er enige om at en av de mest vanlige årsakene til at SMB-er velger å investere i BI, er et behov for automatisering av manuelle prosesser. Utdaterte manuelle prosesser tar opp mye av tiden til ansatte som er ansvarlige for de ulike prosessene, som kan være utvikling av rapporter,

innsending av data, og distribuering av informasjon. Noe data kan ha kritisk leveringstid, og skal brukes videre av andre prosesser. Det blir derfor ønskelig å automatisere slike manuelle prosesser, for å effektivisere tiden til de ansatte, og for å få en raskere levering av data til andre kritiske prosesser. Dette gjøres ved å benytte rapporteringsfunksjonaliteten som tilbys av de fleste BI-plattformene på markedet.

Ila. feltstudiet kom det fram at hovedmotivasjonen til Acando var automatisering av en rekke rapporter. Administrasjonen hos Acando brukte mange timer hver uke på å generere disse rapportene, som skulle leveres både til vanlige ansatte og til ledelsen. Ved å automatisere disse ønsket Acando å oppnå raskere og enklere levering av informasjon til de ansatte, samtidig som tiden til de rapport-ansvarlige i administrasjonen ville bli effektivisert. Dette skulle de oppnå ved å skape en selvbetjeningsmodell, hvor ansatte kunne få rapporter automatisk når de ønsket det. Automatisert rapportering var altså behovet som var avgjørende for Acandos beslutning om å implementere BI, slik som Gartner og Wise mener at det er for mange SMB-er som investerer i BI.

Resultatene fra spørreundersøkelsen viste at funksjonaliteten som var mest brukt av SMB-er var rapportering. Om lag 50 % av alle SMB-er som investerte i BI benyttet rapportering. Ca 30 % benyttet dashboards, som også er en type rapportering. Som motivasjonsfaktorer for bruk av BI, var SMB mest motiverte av raskere datalevering (72 %), operasjonell effektivitet (67 %) og å oppnå en selvbetjeningsmodell (61 %). Det er basert på disse resultatene, rimelig å tro at de fleste SMB-er benytter seg av rapportering, enten ved bruk av vanlige dokument-rapporter, eller dynamiske dashboard-rapporter. Intervjuobjekt C tror at det er vanlig for SMB å først investere i BI pga. rapportering, og senere utvide til flere funksjonelle områder.

8.1.2 Beslutningsstøtte og strategi

I følge All [3] produserer SMB-er ofte massive mengder informasjon, men tenderer til å mangle analytiske kapabiliteter, noe som hindrer de å gjøre forretningsavgjørelser. All tror at mange motiveres til bruk av BI på bakgrunn av dette, og at behovet stadig vokser. Ved hjelp av beslutningsstøtte ønsker bedrifter å få innsikt i dataen på nye måter, og på den måten bedre vite hvilke forretningsvalg som er kloke å ta. Ved bruk av BI kan beslutningsstøtte oppnås ved for eksempel enkel drill-down analyse, eller ved hjelp av mer avansert data mining.

Resultatene fra spørreundersøkelsen viser at om lag 50% av SMB-er motiveres til å investere i BI på grunn av et behov for beslutningsstøtte. Til

sammenligning motiveres 100 % av de store bedriftene av dette behovet. Disse resultatene viser helt klare tegn på at store bedrifter er mer motiverte av beslutningsstøtte, i forhold til de mindre bedriftene. For beslutningsstøtte brukes ofte dashboards for å overvåke bedrifts-metrikker, og data mining for å overvåke mønster i dataen. Blant SMB benyttet om lag 30 % dashboards og 10 % data mining. Blant de store bedriftene brukte 43 % dashboards og 32 % data mining. Dette støtter opp antagelsen om at SMB i mindre grad er motivert av beslutningsstøtte, siden både data mining og dashboards er viktige hjelpemiddel for å gjøre beslutningsstøtte mulig.

I overkant av 40 % av SMB-er- og 80 % av store bedrifter, motiveres til bruk av BI for å få hjelp til bedriftsstrategi. Dette er ikke overraskende, da store bedrifter har et større datagrunnlag, og dermed har mer å lære fra dette. Strategi er mest vanlig å benytte ved å analysere bruksmønstre, og føre strategi deretter, i følge Watsom og Wixom [71].

8.1.3 Analyse

I underkant av 40 % av SMB-ene rapporterte at de brukte drill-down analyser, sammenlignet med over 90 % hos de store bedriftene. Om lag 10 % av SMB-ene svarte at de benyttet seg av data mining-analyser, mens ca 40 % av de store bedriftene benyttet seg av slik analyse-funksjonalitet.

Intervjuobjekt B mener at SMB ikke benytter seg av data mining, fordi de mangler det historiske datagrunnlaget som må ligge i grunnen for å benytte slik funksjonalitet. De sliter også med dårlig datakvalitet, og dårlig struktur på den dataen som finnes. Denne påstanden støttes opp av intervjuobjekt A. Et annet problem relatert til data mining, er mangel på kompetanse. Intervjuobjekt A og B peker begge på det faktum at det er lite data mining-kompetanse i Norge, og da spesielt hos SMB-er. Data mining er også kjent for å være dyrt å gjennomføre, forteller intervjuobjekt A. Han tror at mange velger å ikke gjennomføre data mining, fordi de tror de får større ROI av å gjøre andre analyser, som drill-downs, for å se etter mønster i dataen.

Finansiering, kompetanse, datagrunnlag, og datakvalitet, er alle faktorer som må være på plass for å kunne benytte data mining. Men selv om disse ikke er til stede, og bedriften lar være å bruke data mining, betyr ikke dette at bedriften ikke har et *behov* for å benytte data mining. Det kan fortsatt være tilfelle at bedriften trenger å gjøre data mining-analyser på datagrunnlaget.

Intervjuobjekt B mener at data mining som oftest benyttes for å se på mønster i kundebasen som ligger i CRM-systemer, noe som de færreste SMB-

er har et behov for. Intervjuobjekt A hevder at datakvaliteten til dataen i CRM-systemer ofte er svært dårlig, og at de som ønsker å gjøre data mining på slike systemer har store utfordringer. Dersom dette er rett, er det ikke vanskelig å se for seg at de SMB-ene som faktisk ønsker å benytte seg av data mining for å analysere kundebasene sine, rett og slett ikke er i stand til å gjøre dette.

Resultatene i dette studiet er ikke konkrete nok til å kunne si noe om SMB-s behov for analyser og data mining. Det eneste resultatene viser er at SMB i mindre grad benytter seg av denne funksjonaliteten, sammenlignet med de store bedriftene.

8.1.4 Et helhetlig bilde

Spørreundersøkelsen viste at SMB-er var mindre motiverte for bruk av BI ved samtlige alternativer for funksjonalitet (rapportering, dashboards, drill-down, og data mining). Dette kan bety at SMB-er generelt er mindre motiverte for bruk av BI, og/eller at SMB-er er motiverte av færre faktorer. Dette kan komme av at SMB har mindre behov for BI, og/eller at de setter mindre krav til hva de ønsker ut av BI. Wise [76] tror at SMB-er ofte ikke vet hvilke muligheter som finnes innenfor BI, noe som fører til at de går inn i BI-markedet med lave ambisjoner i forhold til hva de faktisk kan oppnå. Undersøkelsen gjort av BARC [5] i 2010, viste at SMB-er var mer fornøyde med sine BI-implementasjoner, sammenlignet med de store bedriftene. Det ble da foreslått at dette kunne komme av at SMB ikke hadde spesielt store ambisjoner i utgangspunktet. Resultatene fra undersøkelsen gjort i dette studiet støtter både Wise [76] og BARC [5] sine antagelser.

SMB-er benytter seg mest av rapportering, og intervjuobjekt C og Gartner [27] er enige om at rapportering typisk er inngangsporten for SMB til BI. Men selv om SMB-er går inn med små ambisjoner, for eksempel enkel rapportering, er det ikke gitt at de ikke vil ønske å benytte seg av mer avanserte funksjonaliteter senere. Etter hvert som bedriften vokser, vil også behovene og motivasjonene vokse. Det er derfor ikke nok å se på motivasjonene for investering av BI når man undersøker SMB's behov.

8.2 Utfordringer for SMB

8.2.1 Høy Risiko

I følge Fisher [20] har SMB-er mindre rom for feil, sammenlignet med de store selskapene. Dette kommer av at SMB-er har lavere omsetning enn de

store bedriftene, og derfor har mer å tape på å investere i noe som ikke lønner seg. Dermed er SMB-er mer avhengige av å vite at en BI-implementasjon vil lønne seg, før de tar steget til å investere. Utregning av ROI blir da kritisk, og denne er som regel svært vanskelig å regne ut. For å kunne si noe om ROI av en BI-implementasjon, må man blant annet vite følgende:

- Hvor mye kapital den nødvendige programvaren vil koste
- Estimat over implementasjonstiden for BI-løsningen
- Hva man får igjen for implementasjonen i ren bedriftsverdi
- Estimat på hvor mye support som vil bli nødvendig
- Framtidige muligheter for utvidelse

Kostnader for programvare er det enkleste estimatet, da de fleste leverandører gir en god oversikt over hva man får for pengene. Å regne et estimat for implementasjonstid er ikke lett, da det avhenger av hvor raskt utviklerne jobber og hvor mye som skal gjøres. Dette estimatet blir ofte undervurdert, og det er her bedrifter sprenger budsjettet, mener intervjuobjekt C. De som setter dette estimatet må ha god kunnskap om BI, noe som vil diskuteres nærmere i Seksjon 8.2.2..

8.2.2 Mangel på kunnskap om BI

For å regne ut estimat for hvor lang tid en BI-implementasjon vil ta, må man ha dyp forståelse for BI som teknologi. Wise [76] mener at SMB-er ofte har mangelfull kunnskap om hva en BI-implementasjon innebærer, noe som fører til at gapet blir for stort mellom hva som ønskes oppnådd, og det som faktisk må gjøres for å oppnå dette. Man får da et dårlig utgangspunkt for implementasjonen, noe som kan resultere i en BI-implementasjon som blir langt dyrere enn antatt. Et godt eksempel på dette, er Acandos utgangspunkt for sin BI-implementasjon. Veilederne som hadde startet prosjektet var ikke klar over at hele ETL-steget måtte gjennomføres før rapporteringen kunne begynne, noe som førte til at utviklerne sprengte budsjetter allerede i første utviklingsprint. Dette ville vært unngått dersom de som lagde oppdraget hadde hatt grunnleggende BI-kunnskaper.

Ikke alle bedrifter vet hva de ønsker å få ut av BI på sikt, ettersom BI er et område med svært mange muligheter. Resultater har vist at SMB-er ofte går inn i BI-markedet med lave ambisjoner, da typisk med ønske om å benytte rapportering og dashboards. Dersom disse senere gjør mer ut

av BI, vil også ROI kanskje stige. Men grunnet de lave ambisjonene som utgangspunkt, vil ikke disse vinningene tas hensyn til i utregning av hva bedriften får igjen i bedriftsverdi. Dette kan føre til at bedrifter regner ut feil ROI, og kanskje velger å ikke investere i BI, som resultat av de ikke visste hva mulighetene var, og/eller hadde for lave ambisjoner.

Wise [76] mener at SMB-er ofte ikke har god nok kompetanse om BI til å gjøre en komplett BI-implementasjon. Bedriftene har teknologien og interaktiviteten tilgjengelig, men opplever et pedagogisk sprik mellom teknologien og de ansatte. Bedriftene kan da velge å utdanne sine ansatte, men blir da avhengig av å bevare kompetansen innad i bedriften. Gartner [27] mener at man i slike tilfeller er nødt til å leie inn ekstern BI-kompetanse. I Acandos tilfelle gikk det greit å la utviklerne lære seg BI på egen hånd, men utviklingen gikk treigt, som resultat av at utviklerne ikke kunne teknologien på forhånd. Acando mistet også kompetansen da sommerjobben tok slutt, noe som sannsynligvis øker terskelen for at Acando vil fortsette utviklingen.

8.2.3 Ansatte

Gartner [27] påpeker at det ofte er en utfordring å få samarbeidsvilje fra de ansatte, i starten av BI-prosjekter. Ansatte er skeptiske til å automatisere manuelle prosesser, da dette vil bety at de mister denne oppgaven. Mange føler også eierskap til data, og har ikke noe ønske om å dele denne. Under implementasjonen hos Acando, fikk utviklerne inntrykk av at de ansatte mer enn gjerne ville slippe å gjøre de manuelle prosessene, og at de var svært villige til å samarbeide. Dette kan komme av at de visste at de ville beholde jobbene sine uansett, og at de heller fikk bruke tiden sin på mer nyttige og spennende arbeidsoppgaver. Det kan tenkes at motvilje fra de ansatte er et mindre problem i land som Norge, hvor det ikke er like lett for arbeidsgiver å si opp de ansatte.

Gartner [27] mener også at kniving om budsjett er et vanlig problem for SMB-er som ønsker å implementere BI. Det kan være uenigheter om hvilke investeringer som skal prioriteres, da ulike avdelinger har ulike ønsker. Gartner mener at man må få de ansatte til å samarbeide, for å oppnå et felles solid datagrunnlag, i stede for å sitte på litt data hver, uvillige til å dele. For å få til dette må man klare å overbevise de ansatte om at BI er en god ide, som vil komme alle til gode.

8.2.4 Inkonsistente bedriftsmetrikker

Gartner [27] påpeker at SMB-er har en stor fordel med at ledelsen ikke er like stor som hos de store bedriftene. Færre i ledelsen betyr færre involverte parter i byråkratiske avgjørelser. Dette gir muligheten til rask respons på tilstandsendringer i bedriften, og bedriften vil typisk være smidige. Men denne smidigheten er ikke gratis; bedriften endres kanskje så raskt, og så ofte, at det setter opprettholdelsen av konsistente og reliable bedriftsmetrikker på spill. Det kan bli vanskelig å lage standardiserte, konsistente målinger til datagrunnlaget, når dataen hele tiden endres. Uten konsistent, pålitelig bedriftsinformasjon, og gode analytiske prosesser, er det umulig for bedrifter å gjøre bedriftsavgjørelser basert på BI.

8.3 Bruk av OSBI

8.3.1 Fordeler

En av årsakene til at OS-programvare har blitt populært, er at det gir brukeren mye frihet til hvordan han ønsker å benytte programvaren. Denne friheten er svært tydelig i OSBI-markedet, hvor leverandørene samarbeider for å gi brukeren mulighet til å velge mellom verktøy. I feltstudiet ble det klart at Jaspersoft og Pentaho har hatt et tett samarbeid under utviklingen av ulike verktøy, og at de bevisst har brukt de samme filformatene. Slik er det ikke for de tradisjonelle plattformene, og fleksibilitet og modularitet kan derfor sies å være en av de største fordelene ved bruk av OSBI. Leverandøruavhengighet er et resultat av modulariteten og fleksibiliteten, og er også stor fordel ved å benytte OSBI.

Lav pris på programvaren er en av de mest avgjørende faktorene for at bedrifter velger OSBI framfor kommersiell BI. I spørreundersøkelsen kom det fram at 90% av de som valgte OSBI-programvare gjorde dette valget delvis på grund av lav pris. Dette var den desidert mest motiverende faktoren for valg av plattform, etterfulgt av funksjonalitet, som motiverte om lag 70% av OSBI-brukerne.

Ved å benytte OS-programvare får man mulighet til å delta i utviklingen av selve programvaren. Dersom man for eksempel har oppdaget en bug som man gjerne skulle ha fikset til neste release, kan man fikse denne selv. I tillegg er det mulig å komme med innovative innspill til utviklingen, og dermed styre den evolusjonære retningen som programvaren tar.

8.3.2 utfordringer

Kompetansen på OSBI er generelt lav i Norge, noe som gir utfordringer både for utvikling av- og vedlikehold av OSBI-implementasjoner. Utviklere må enten være ansatte i bedriften som har tilstrekkelig kompetanse, eller leid inn fra konsulent-selskaper. Få norske bedrifter spesialiserte seg på OSBI, noe som gjør kompetansen vanskelig å finne. Og dersom man finner kompetansen, så er man avhengig av å ha denne tilgjengelig i framtiden, for vedlikehold og/eller videreutvikling av løsningen. For å slippe problematikken rundt ekstern anskaffelse av kompetanse, kan bedriften gjøre som Acando gjorde, og la utviklerne lære seg BI og BI-verktøyene på egen hånd. Man risikerer da at kompetansen forsvinner dersom utviklerne slutter i bedriften, slik som også var tilfelle hos Acando.

En av de største risikoene ved bruk av OS-programvare er faren for at prosjektet legges ned. Utviklingsmiljø som drives av frivillige utviklere er avhengig av at utviklere fortsatt er motiverte til å jobbe på prosjektet. Dersom de ikke lenger får noe ut av det, er terskelen svært lav for å slutte på prosjektet, og det er derfor veldig vanlig at OS-prosjekter dør ut. Hos OSBI-leverandører som Pentaho og Jaspersoft er det lite trolig at dette vil skje med det første, ettersom de har mange ansatte som jobber på prosjektet på heltid, og dermed ikke er like avhengige av frivillighetsarbeid. Men bruk av OS-programvare er forbundet med en risiko i industrien, og mange ønsker derfor å ta avstand fra det. Intervjuobjekt A fortalte at hans bedrift holder seg unna OSBI, fordi de ønsker å vise sine kunder at de tar utviklingen på alvor, og ikke "tar sjansen" på å bruke OS.

En fare ved bruk av OSBI er missoppfatningen av at billig programvare tilsvarer billig utvikling; arbeidstimer for utvikling er fortsatt en stor utgift, enten man benytter egne ansatte eller eksterne konsulenter. Intervjuobjekt A mente at man ved bruk av OSBI kanskje ville bruke langt flere timer på utvikling, siden verktøyene ikke er like brukervennlige. Mangel på opplæring, support og dokumentasjon fører også til flere timer brukt på utvikling. Under utviklingen av Acandos BI-implementasjon ble det flere ganger komplett stopp i framgangen på prosjektet, fordi man ikke skjønnte feilmeldingene eller fordi noe var feil uten noe melding i det hele tatt. I slike tilfeller ville det vært svært nyttig å ha support tilgjengelig, hvor man da kunne beskrevet problemet, og fått råd og anbefalinger uten å måtte lete selv. Utviklingen hos Acando var preget av leting på forum og rådføring på chate-tjenester. Spørreundersøkelsen viste at OSBI-utviklere brukte BI-forum mer enn dobbelt så mye som utviklere som brukte kommersiell BI. De som utviklet med kommersiell BI støttet seg mer på kollegaer, og er-

faringer fra tidligere prosjekter. Det er rimelig å anta at det tar lengre tid å lete etter hjelp på forum, enn det tar å ringe inn til support eller lete i dokumentasjonen. Det kan derfor se ut som at OSBI-utvikling faktisk krever flere arbeidstimer, og at dette burde tas hensyn til når bedriften vurderer om de skal benytte OSBI.

Gartner [27] mener at bedrifter aldri må basere valg av OSBI-programvare utelukkende på grunn av lav pris, men at de må bruke de samme beslutningskriteriene som de ville benyttet for valg av proprietær BI. På denne måten unngår de å velge en programvare som egentlig ikke passer bedriften. IT Novum [32] er enig i at BI-plattformen må passe bedriftens behov, for å skape et best mulig utgangspunkt for implementasjonen. Dersom plattformen for eksempel ikke tilbyr den funksjonaliteten, supporten, eller sikkerheten som bedriften har behov for, kan dette føre til at implementasjonen blir mislykket.

8.3.3 Modenhhet

OSBI har gått fra å bestå av umodne enkeltstående verktøy på starten av 2000-tallet, til å i dag bestå av hele plattformer, med alle nødvendige verktøy for å gjøre en full BI-implementasjon. Man kan velge mellom helt åpne versjoner, og kommersielle versjoner av disse plattformene, avhengig av hva man ønsker.

Gartner [7], Thomsen og Pedersen [66], IT Novum [32] og Golfarelli [28] er enige om at Pentaho er den mest lovende plattformen blant OSBI-alternativene. Denne plattformen ble benyttet på implementasjonen til Acan-do, hvor feltstudiet til dette studiet ble gjennomført.

I 2008 gjennomførte Thomsen og Pedersen [66] en test hvor de vurderte modenheten til enkeltstående verktøy. De brukte da kategoriene DBMS, ETL, OLAP-server og OLAP-klient. De lagde en liste over kriterier som skulle vurderes, og vurderte hvert verktøy basert på disse. Pentaho kom totalt sett best ut i denne testen, og de mente at Pentaho hadde gjennomgått en bratt modningskurve de siste 4 årene. Det var inkludert en rekke nye verktøy i plattformen, og plattformen var generelt kraftigere, mer brukervennlig, og nærmere et godt alternativ til de tradisjonelle BI-plattformene.

I vurderingen til Thomsen og Pedersen [66] kom det fram at ETL-verktøyet hadde et godt brukergrensesnitt med mye funksjonalitet. Det eneste de savnet var mulighet for å kjøre parallelle jobber, for å kunne forbedre prosessytelse. Dette har i ettertid blitt implementert i ETL-verktøyet, noe som vil ha mest å si for implementasjoner som har tidskritiske prosesser som kjører, hvor en seriell jobb-prosessering ikke er tilstrekkelig. Dette er

typisk prosesser som brukes av bedrifter som gjør real-time analyser til dashboards eller vanlige rapporter. Resultatene fra spørreundersøkelsen viser at store bedrifter har langt høyere fokus på rask informasjonslevering og operasjonell effektivitet, som er ytelsesrelaterte områder. Det er derfor mulig at parallellprosessering kan få mest å si for de større bedriftene som ønsker å benytte Pentaho. Ellers rapporterte Thomsen og Pedersen [66] at de savnet brukergrensesnitt for bygging av kube, da disse måtte utvikles manuelt i XML-filer. Et slikt GUI for kubebygging har blitt tilgjengelig på plattformen i ettertid, og under feltstudiet ble dette verktøyet beskrevet som noe umodent, men relativt enkelt å bruke. De ulike verktøyene var tilsynelatende av svært varierende kvalitet, og det var ikke vanskelig å se hvilke verktøy som hadde eksistert lengst. De verktøyene som ble brukt av flere prosjekter, utmerket seg positivt.

8.3.4 Åpen OSBI eller kommersiell OSBI?

Pentahos funksjonalitet og arkitektur ble testet av Golfarelli [28] i 2009, sammen med flere andre plattformer. Golfarelli kartla på den tiden hva den kommersielle Pentaho-plattformen kunne tilby, som ikke fantes i den helt åpne versjonen:

- Bedre administrative konsoller
- Lettere konfigurering av systemfiler
- Prosessovervåkning
- Debuging-GUI for ETL
- Større dokumentasjon, samt en kunnskapsbase

Etter at Golfarelli hadde gjort testen, konkluderte han med at den helt åpne versjonen ikke burde brukes når man skal utvikle kritiske applikasjoner som skal brukes i komplekse organisasjoner. Dette begrunnet han med at det ved hjelp av den åpne versjonen ikke var lett å administrere løsningen og å overvåke prosesser som kjørte. Den kommersielle løsningen, derimot, mente han var svært godt egnet til dette, da den hadde utvidet administrative funksjonaliteter, samt prosessovervåking. Ila. feltstudiet på Acando, ble det klart at den kommersielle plattformen fortsatt har disse funksjonelle fordelene. Det eksisterer nå også en dashboard-designer som kun er tilgjengelig i den kommersielle versjonen. Jeg og min kollega fikk erfare mye av det Golfarelli beskriver, med vanskeligheter for administrering av systemet, da spesielt med tanke på synkronisering og oppdatering av de ulike dataene. Det er

svært ugunstig at programvare kommer i veien for dagligdagse operasjoner, i dette tilfelle administrering, noe som svekker kvaliteten og brukbarheten av programvaren. Bedrifter som er avhengig av å ha god administrasjonsfunksjonalitet, prosessovervåking og support, har derfor ikke nok med den helt åpne OSBI-plattformen, da denne mangler disse tilbudene. Den mangler også et grensesnitt for å designe dashboards, som krever kjøp av den kommersielle versjonen. De som er usikre på om de trenger den funksjonaliteten som kun tilbys av den kommersielle versjonen, har mulighet til å starte med å benytte den helt åpne versjonen, og heller gå over til enterprise-versjonen etter hvert.

Intervjuobjekt C uttalte at han ikke trodde at det fantes noen i Norge som hadde utviklet en BI-implementasjon ved å bruke en helt åpen OSBI-plattform. Han hadde selv jobbet med OSBI-utvikling i mange år, ved å bruke Pentaho Enterprise-versjonen, og mente at denne versjonen var nødvendig for å kunne gjøre en god implementasjon. I spørreundersøkelsen kom det fram at 7 av de 10 respondentene med OSBI-erfaring hadde benyttet enterprise-versjonen. De resterende 3 respondentene hadde alle jobbet på Acandos BI-implementasjon, noe som kan tyde på at Intervjuobjekt C har rett, og at Acando var et spesielt tilfelle. Jeg og min kollega var enige om at det hadde vært en fordel for oss, som utviklere, om Acando investerte i enterprise-versjonen, da utviklingen gikk mye langsommere med den åpne versjonen. Acandos implementasjon var relativt liten, og det er rimelig å anta at å bruke den åpne versjonen på en større implementasjon ville blitt for omfattende. Det vil komme til et punkt hvor økte kostnader for utviklingstimer overstiger kostnadene spart på programvare.

8.3.5 OSBI for SMB

Golfarelli [28] mente at OSBI helt klart ikke er like sofistikert som de store kommersielle plattformene, men at OSBI likevel må ansees som et valid alternativ. Spesielt for SMB-er som ofte ikke har kritisk datakvalitet og ytelse, er OSBI et godt alternativ. I spørreundersøkelsen svarte 10 av 10 respondenter med OSBI-erfaring at OSBI kan være et godt alternativ for SMB-er, mens 6 av 10 mente at OSBI kan være et godt alternativ for store bedrifter. Det betyr at 40 % av respondentene som hadde erfaring med OSBI *ikke* mente at OSBI kunne være et godt alternativ for store bedrifter. Det var 100 % enighet om at OSBI kan være et godt alternativ for SMB, men *kan være* betyr ikke at det alltid vil være det. Ulike SMB-er har ulike behov. Intervjuobjekt C mener OSBI er godt nok for enkelte SMB-er, men ikke alle. Som forutsetninger for at en bedrift skulle lykkes med OSBI satte han intern

entusiasme- og kultur for bruk av OS-programvare i bedriften. Dersom de ansatte er skeptiske til bruk av OS, vil dette kunne skape dårlig stemning, som igjen kan være ødeleggende for utviklingen. Intervjuobjekt C mente videre at det er helt essensielt å ha tilgang på kompetanse, og at denne er vanskelig å finne for OSBI-plattformer. Disse forutsetningene gjelder selv om man velger enterprise-versjonen av OSBI-plattformen.

Intervjuobjekt A mener at OSBI-plattformene er svært dårlige til å markedsføre seg selv, og at mange derfor ikke har hørt om de. De kommersielle, derimot, er flinke til å markedsføre seg, og får derfor mer oppmerksomhet. Han beskrev det norske BI-markedet med en isfjells-analogi; toppen av isfjellet er det du ser, og det er her du finner de kommersielle BI-alternativene, som er gode til å synliggjøre seg selv på markedet. Dersom man ikke vet at OSBI-alternativene finnes, kan man heller ikke benytte de, med mindre man aktivt oppsøker de. De som typisk oppsøker KOSBI er de som ikke har råd til standard kommersiell programvare i utgangspunktet, eller de som foretrekker OS-programvare. Blant disse finnes typisk SMB-er, forskere [28], og offentlig sektor. Golfarelli mener at forskere bruker BI for å prototype, teste, og analysere datagrunnlag. Han mener også at mange offentlige instanser foretrekker OS-programvare, og at OSBI derfor kan bli en viktig del av IT for offentlig sektor.

8.3.6 Framtidsutsikter

Framveksten av OSBI har skjedd fort, og OSBI har gått igjennom den samme prosessen som kommersiell BI, men de begynte denne prosessen 20 år senere. Pentaho og Jaspersoft har gjennomgått en kommersialiseringsfase, hvor de gikk fra å kun utvikle helt åpen programvare, til også å tilby tjenester med fortjeneste. I spørreundersøkelsen kom det fram at 9 av 10 respondenter med OSBI-erfaring, mente at OSBI-bedriftsmodellen er levedyktig på lang sikt. Golfarelli [28] mener at OSBI har vært på markedet så kort at det er umulig å si om plattformene vil tjene nok fra de tjeneste som de tilbyr i enterprise-versjonene, til å bli på markedet.

Gartner [7] spådde i 2008 at OSBI ville være mest brukt av SMB-markedet fra 2008 til 2012, og at OSBI deretter ville være et helt vanlig valg på OSBI-markedet. Dersom Gartner har rett, burde det være mulig å se klare endringer i BI-markedet. Det har skjedd en stor vekst innen OS-BI fra 2008 til i dag, både på modenhet av OSBI-verktøy, popularitet av verktøyene, og antall brukere. Men selv om populariteten av verktøyene har økt, er det fortsatt mange som ikke kjenner til disse alternativene. Av konsulentene som svarte på spørreundersøkelsen gjort i dette studiet, var

kjennskapet til de ulike OSBI-plattformene overraskende lav. Dette var til tross for at respondentene hadde mye erfaring innenfor BI. Av de som hadde jobbet lengre enn 15 år i BI-bransjen, hadde i underkant av 50% hørt om Pentaho, mens om lag 65% hadde hørt om Jaspersoft. Det er tydelig at disse OSBI-plattformene fortsatt ikke er blitt helt vanlige valg, ettersom BI-konsulenter med så mye som 15 års erfaring ikke har hørt om disse alternativene. Gartner vil derfor trolig få feil i antagelsen om OSBI's status i 2012.

Intervjuobjekt A tror OSBI kan bli mer brukt i framtiden, dersom kommersiell BI ikke går ytterligere ned i pris. Han peket på det faktum at kommersiell BI i dag koster en femtedel av hva det kostet for 5 år siden, og at det forventes at prisene vil gå ned ytterligere. OSBI kan bidra til å presse prisene, men dersom prisene ikke går ned ytterligere, vil kanskje OSBI vinne en større del av markedet.

8.4 Anbefalt tilnærming til BI for SMB

8.4.1 Valg av programvare

IT Novum [32] og Gartner [7] er enige om at det er viktig at BI-plattformen som velges, passer bedriftens behov. Vanlige beslutningskriterier når man skal velge programvare er hvordan funksjonalitet man ønsker, ønsket arkitektur, grad av sikkerhet, modularitet i infrastrukturen, hvor lett det er å skaffe kompetanse fra konsulenter, og support-tjenester fra leverandør. Noen SMB-er velger plattform etter pris, for å få lavere kostnader på BI-prosjektet. En plattform burde aldri velges utelukkende på grunn av pris, da dette kan føre til at bedriften velger en plattform som ikke passer bedriftens behov. Bedrifter som vurderer å benytte OSBI, burde bruke de samme beslutningskriteriene som de ville brukt for tradisjonell BI. På denne måten unngår man å bli blendet av den lave prisen. Intervjuobjekt A mener at bedriften har et feil utgangspunkt, dersom for dyr programvare er et problem. Utgifter for programvare er små, sammenlignet med utgifter på konsulenttimer. Dersom man først har konstatert at man behøver BI, burde man kjøpe en trygg programvare, mener han.

Det er i dette studiet testet bruken av OSBI-plattformen Pentaho, som av flere [7, 28, 32] omtales som den mest lovende plattformen på markedet. Plattformen har hatt stor vekst de siste årene, og viste seg å være svært moden. Framtidsutsiktene lover bra, og det er lite trolig at prosjektet vil legges dødt i nærmeste framtid. Hvorvidt en slik OSBI-plattform burde benyttes av SMB-er, avhenger av bedriftens behov. Dersom bedriften skal benytte

BI som en intern løsning, som ikke har et behov for dashboard-designing, god administrasjonsfunksjonalitet og prosessovervåking, kan den helt åpne OSBI-versjonen vurderes å brukes. Men om noen av disse behovene eksisterer, burde bedriften se på kommersiell OSBI eller tradisjonell BI i stede. Dersom man ønsker å velge OSBI, enten helt åpen eller kommersiell, må man forsikre seg om at man har tilgjengelig kompetanse om plattformen, enten fra eksterne konsulentfirmaer eller fra egne ansatte. Man kan velge å gi ansatte opplæring, men da blir disse ansatte svært verdifulle for bedriften, dersom det er snakk om få personer. Det er også svært viktig å være klar over at billig programvare kan føre til større kostnader enn ved bruk av kommersiell programvare, da man risikerer å bruke lengre tid på selve utviklingen.

8.4.2 Håndtering av risiko

En av de største utfordringene til SMB-er, er risikoen rundt investering i BI. Risikoen er knyttet til begrensede ressurser tilgjengelig. Det er derfor viktig for SMB-er å håndtere denne risikoen, ved å estimere ROI før man bestemmer seg for å gjøre en BI-implementasjon. Dersom bedriften er i stand til å gjøre en god beregning av ROI, vil den få en god pekepinn på hvorvidt de burde gjennomføre implementasjonen eller ikke. En slik risikovurdering anbefales derfor å gjøres grundig. I dette regnestykket inkluderer man kostnader for programvare, et estimat for implementasjonstiden, hvor mye man vinner i ren bedriftsverdi, og hvor mye drift av systemet vil koste.

Gartner [27] foreslår at ROI skal beregnes for hver enkelt prosess som skal automatiseres eller lages ny. Følgende må da vurderes:

- Prosessens viktighet for bedriften som helhet
- Ressurser involvert (innsats som kreves for dataforberedelser)
- Hvordan dataen/informasjonen vil konsumeres videre av andre prosesser og analyser [27].

Det er svært viktig at de som regner ut ROI for hver prosess har god innsikt i både forretningssiden og den tekniske siden av implementasjonen.

8.4.3 Sikre kunnskap og kompetanse

For å være i stand til å beregne korrekt ROI, må man ha bred kunnskap om både BI som teknologi, og bedriften som skal implementere det. Gartner [27] og Wise [76] er enige om at det må fokuseres på kunnskapsnivåene til alle de involverte partene i utviklingen. Dersom denne er for lav for de ansatte i

bedriften, må kompetanse skaffes utenifra. Gartner [27] foreslår at man kan sette sammen et utviklingsteam bestående av en kombinasjon av interne og eksterne ressurser. Ekspertise i analyse, forretning og intern arkitektur burde ideelt komme fra interne ressurser, mens teknisk kompetanse kan leies fra eksterne konsulentselskaper.

Dersom bedriften ikke har råd til ekstern utviklingshjelp, må egne ansatte læres opp, for å besitte den kunnskapen som trengs. Man må da ta hensyn til at BI-løsningen skal være mulig å drifte, selv når de/den ansatte som gjorde utviklingen slutter i bedriften. Dette kan gjøres ved å fokusere på god dokumentasjon på alt som gjøres, samt kompetanseoverføring mellom ansatte. Ved å alltid ha minst to ansatte med den samme kunnskapen om BI-løsningen, vil ikke bedriften risikere å få problemer dersom noe skulle skje med en av de.

8.4.4 Skape et realistisk utgangspunkt

Bedriften som implementerer BI må være klar over hvordan implementasjonen vil se ut, for å unngå å sprengte budsjettet. Ved å gjøre et godt ROI-estimat, og skaffe kunnskap/kompetanse om BI, vil man være i stand til å skape et realistisk bilde av implementasjonen før utviklingen starter. Det er også en stor fordel å vite hvilke muligheter som finnes innenfor BI, slik at man kan tilrettelegge for disse mulighetene. Gartner [27] mener at SMB-ene trenger en tilnærming til BI hvor de tenker stort, men starter smått, for å holde fokus på en solid grunnmur. Man kan for eksempel starte ved å fokusere på rapportering, og senere ekspandere til analyser og andre funksjonelle områder. Ved å se på muligheter og begrensninger, blir bedriften godt rustet til å skape et realistisk bilde av implementasjonsprosessen, samtidig som de ser hvordan implementasjonen kan utvides etter hvert.

8.4.5 Kommunikasjon

Når man skal implementere en BI-løsning, er det viktig at de ansatte i bedriften er samarbeidsvillige. Det er disse som har kunnskap om prosessene som skal automatisere, og det er ofte disse som kjenner datagrunnlagene best. Gartner [27] mener at det burde holdes en proaktiv kommunikasjon med de ansatte involvert, for å justere forventningene deres, og forberede de på at innovative endringer som vil komme. Dette gjelder både ansatte som vil miste arbeidsoppgaver (som følge av automatisering), og ansatte som i større grad må dele data fra sine systemer. På denne måten får de ansatte mulighet til å gradvis forberedes på et samarbeid, slik at de kanskje er mer

positivt innstilte når tiden for endringene kommer.

Gartner[27] foreslår at man her må gjøre en top-down-tilnærming for å få en god infrastruktur for hele bedriften, og da ha fokus på å holde sentrerte og konsistente bedriftsstandarder. Bedriftskulturen må endres fra å la ulike enheter/avdelinger holde data i siloer, og heller oppfordre disse til å samarbeide mot et felles datagrunnlag og et felles mål. En mulig løsning er å la minst én representant med ansvar for hver datakilde delta i konstruksjonen av infrastrukturen. Ved å samarbeide på en slik måte vil man unngå at enheter i samme bedrift kniver om budsjettet, og heller forsøker å oppnå langsiktig bedriftsverdi, ved å investere i en felles løsning for datalagring.

Intervjuobjekt C påpeker at det er spesielt viktig med kommunikasjon mellom forretningssiden og den tekniske siden under BI-implementasjonen. Ettersom BI er forretningsdrevet, må forretningssiden inkluderes mer i utviklingen, sammenlignet med tradisjonelle systemutviklingsprosjekter. Man kan for eksempel starte med å fokusere på en gitt KPI, og gjøre analyser og rapporter ferdige for denne, før man beveger seg over på neste. På denne måten oppnår man en iterativ utvikling, som inkluderer forretningssiden for hver eneste iterasjon. Dette kan også bidra til mindre utålmodighet fra de som venter på resultater, og større samarbeidsvilje fra forretningssiden.

8.4.6 Tilrettelegge for vekst

SMB-er har typisk færre personer i ledelsen, og dermed også færre personer inkludert i byråkratiske avgjørelser, sammenlignet med store bedrifter. Dette fører til at slike avgjørelser blir lettere å ta. Dersom en SMB velger å bruke BI for bedriftsstrategi og forretningsstøtte, vil denne lettheten rundt byråkratiske styringsprosesser komme til bedriftens fordel, da det vil bidra til at de raskere kan reagere på tilstandsendringer i bedriften (for eksempel økonomiske stormer, eller framvoksende trender [27]). Denne muligheten til å lett endres, gjør at SMB-er er mer smidige enn store bedrifter.

Gartner [27] og Kernochan [34] er enige om at et viktig kriterium for at SMB-er skal bruke BI optimalt, er utnyttning av denne fleksibiliteten og smidigheten, som ofte eksisterer hos SMB-er. For å få til dette mener Gartner [27] at SMB-er trenger fleksibilitet i informasjonshåndteringen. Han foreslår følgende punkter for å sikre dette:

- Det må dannes klare forståelser om forventningene til levering av informasjon
- Det må lages prioritetsnivå for dataen som skal håndteres

- Det må lages en plan for bedriftsstrategien framover
- Bedriften må være innstilt på å lære av sine feil
- Det må holdes en kontinuerlig kommunikasjon mellom alle involverte partene
- Ledelsen må være åpen for endringer i planen, samt skiftende bedriftsprioriteter

Fleksibiliteten gjelder ikke bare på kort sikt, men også på lang sikt; det må tilrettelegges for inkrementell forretningsbygging og framtidig vekst. En solid grunnmur for BI-implementasjonen, åpner mange muligheter for bedriften. Ved å lage et godt datavarehus, som lagrer både nødvendige data og data som *kan komme til* å bli nyttige etter hvert, vil disse dataene ligge klar dersom nye krav til informasjon etter hvert oppstår. Kernochan [34] mener at nesten alle SMB-er har en “ønskeliste” over ting de ønsker å oppnå på sikt. De burde derfor lage en liste over disse, med prioriter. Dermed vil disse være klare, når bedriften i framtiden har samlet tilstrekkelige mengder historiske data, og dermed kan gjøre ønskelige analyser.

Kapittel 9

Konklusjon

Billigere BI-programvare, og et stadig økende behov for bedriftsinformasjon, har ført til at flere og flere SMB-er våger å nærme seg BI-markedet. Disse står overfor litt andre utfordringer enn de store bedriftene; utfordringer som må håndteres for at bedriftene skal lykkes med bruk av BI.

SMB-er har lavere omsetning, og derfor mindre ressurser å bruke på investering i nye teknologier. Det er derfor viktig å bruke tid på å vurdere ROI for en BI-implementasjon, slik at man unngår å investere i noe som ikke lønner seg. For å beregne ROI kreves god kompetanse om både bedriften og BI som teknologi. Ved å vite hvordan en typisk BI-implementasjon ser ut, unngår man å lage tidsestimater som sprenger budsjettet, men får heller et realistisk utgangspunkt.

Når bedriften skal velge programvare, vil det kanskje friste å velge en billig programvare, eller helt gratis OSBI-programvare. Resultater tyder på at bruk av helt åpen OSBI kan føre til flere timer brukt på utvikling, samtidig som man får dårligere administrasjonsstøtte, og ingen verktøy for dashboard-utvikling og prosessovervåkning. Bedrifter som vurderer å benytte OSBI må passe på at plattformen ikke velges kun fordi den er billig, men at den også passer bedriftens behov, som arkitektur, sikkerhet, funksjonalitet, og støtte for support. Uansett om man velger OSBI eller tradisjonell BI, må man forsikre seg om at kompetanse om den valgte plattformen er tilgjengelig, for utvikling og vedlikehold av løsningen. Dette blir en større utfordring ved bruk av OSBI-plattformer, da det finnes lite kompetanse på disse i Norge. Man kan velge å utdanne egne ansatte for å gjøre utviklingen, men man må da forsikre seg om at tapet av de ansatte ikke vil påvirke levedyktigheten til BI-implementasjonen videre.

For å skape samarbeidsvilje hos de ansatte, burde ledelsen holde en

proaktiv kommunikasjon om BI-implementasjonen, og oppfordre til samarbeid i mellom avdelingene. På denne måten unngår man at ansatte setter seg på bakbena, motvillige til å dele informasjon. Det er spesielt viktig med kommunikasjon mellom forretningssiden og den tekniske siden av implementasjonen, ettersom BI er forretningsdrevet. En iterativ implementasjon, med én KPI i fokus for hver iterasjon, er en mulig løsning for å sikre en god kommunikasjon.

SMB-er er smidige, og i stand til å reagere raskt på tilstandsendringer i bedriften. Dette burde utnyttes, ved at man skaper en fleksibel informasjonshåndtering, og gjør bedriften klar for endringer. På denne måten oppnår bedriften en inkrementell forretningsbygging, som tilrettelegger for framtidig vekst. Bedriften vil da være godt rustet for framtidige endringer i informasjonsbehovene.

Dette studiet har vært preget av lite sammenligningsgrunnlag for innsamlet data. Det ville vært interessant å gjøre flere case/feltstudier på andre BI-implementasjoner for SMB-er, for å se hvordan Acando skiller seg fra andre tilfeller.

Appendiks

Tillegg A

Spørsmål fra spørreundersøkelse

(på engelsk)

A.1 Del 1 - Introduction

Introduction



This survey is conducted by a Master student at NTNU (Norwegian University of Science and Technology).

The purpose of the survey is to collect empirical data regarding the need for- and development of Business Intelligence solutions, and is therefore most relevant for technical BI developers.

The survey will take 5-10 minutes to complete, and your answers will be kept anonymous.

Thank you!

Next >>

A.2 Del 2 - Background

Question 1: How old are you?

- 20-30
- 30-40
- 40-50
- 50-60
- 60 <

Question 2: Are you working at an established Business Intelligence Competency center?

- Yes
- No
- Don't know

Question 3: Which of the following areas have you worked in? (Select all that apply)

- Business Strategy
- Consulting Data
- Integration
- Data warehouse
- Reporting
- Analysis

Question 4: How many BI projects have you worked on?

- 1
- 2-3
- 4-6
- 7-10
- 11-15
- 15 <

Question 5: For how many years have you been working with BI?

- 1
- 2-3
- 4-6
- 7-10
- 11-15
- 15 <

Question 6: How well do you know the following BI-tools?

Her skulle respondentene velge ett av tre valg for hvert verktøy:

- Never heard of
- Heard of
- Worked with

Alternativer:

- Oracle Business Intelligence Suite
- IBM - BI Installations
- SAS Enterprise BI Server
- Jaspersoft
- Cognos Business Intelligence
- Business Objects
- Actuate BIRT
- Pentaho BI Suite
- SAP - Business Warehouse
- Hyperion System BI
- Microsoft BI Systems
- SpagoBI

A.3 Business Intelligence in practice

For the questions on this page, consider the last BI project you worked on / is working on.

Question 7: What is / was your customers primary industry?

- Government/Public sector
- Financial services
- IT and Technology
- Manufacturing
- Professional services
- Consumer goods
- Energy and natural resources
- Telecoms
- Construction and real estate
- Transportation, travel and/ tourism
- Entertainment, media and publishing
- Logistics and distribution
- Education
- Other

Question 8: What was your customers main motivations for investing in a BI implementation? (Skip if you don't know)

Her skulle respondenten krysse av på en skala med følgende alternativer:

- Very important
- Important
- Neutral
- Unimportant

Alternativer:

- To share information with customers, suppliers and partners
- To move users toward a self service model of information delivery
- To decrease business costs and improve operational efficiency
- To speed up and improve the organization's decision making ability
- To better align with and track against corporate strategy
- To better manage a specific operational process in a more timely fashion
- To respond to user needs for data on a timely basis

Question 9: Approximately how many of your customers employees will regularly be using the BI solution?

- 1-5
- 6-10
- 11-25
- 26-50
- 50-100
- 100-200
- 200-500
- 500-2000
- 2000-5000
- 5000 < Don't know

Question 10: What percent of those BI users would you consider to be power users (i.e. have the ability to drill down, analyze data, and create reports) - as opposed to casual users (periodic viewing of static reports)?

- 0%
- 5%
- 10%

- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%
- Don't know

Question 11: What percent of employees with BI access will be using the solution in a self-service, non IT-assisted capacity?

- 0%
- 5%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%
- Don't know

Question 12: Which BI platform(s) was used?

- Oracle Business Intelligence Suite
- IBM - BI Installations
- SAS Enterprise BI Server
- Cognos Business Intelligence
- Business Objects
- SAP - Business Warehouse
- Hyperion System BI
- Actuate
- Microsoft BI Systems
- Other (Specify)

Question 13: Why was this platform chosen? (Choose up to three reasons)

- Ability to support large numbers of concurrent users
- Availability of local support
- Bundled with another product
- Chosen vendor did a better sales job
- Completed "proof of concept" faster or better than others
- Corporate standard
- Ease of use for application builders
- Ease of use for end-users
- Fast performance
- Functionality/ product features
- Integrates with other products already in use
- Large data handling capacity

- Low price
- Product reputation
- Range of server platforms supported
- Vendor relationship/ reputation
- Web architecture
- Other (Specify)

Question 14: In which of the following BI activities does your customer engage? (Select all that apply)

- Dashboard visualization of key metrics
- Regular / periodic report generation and delivery
- Ad hoc query and reporting
- Scoreboards and performance management
- Predictive analytics
- Real-time analytics
- Drill-down analytics
- Don't know

Question 15: Which of the following have you been working with on this project, and what licences did the tools have? (Only check the ones you have worked with)

- Data integration
- Meta data management
- Data warehouse
- Development and delivery of reports
- Development and delivery of BI analytic applications
- Development and delivery of dashboards
- Data modeling development

- Cube building

Question 16: To what extent have you used the following sources, when in need of development support?

Her skulle respondenten vurdere hver kilde på en skala med følgende alternativ:

- Very large
- Large
- Neutral
- Small
- Very Small
- Don't know

Alternativer:

- Documentation for the BI-tools used
- Colleagues
- Forums for the BI-tools used
- Customer support from the BI-tools used
- Examples from previous projects you worked on
- Non-tool specific forums for BI developers

Question 17: Please arrange the usefulness of the sources in the previous question.

Her skulle respondenten vurdere nyttigheten til hver kilde på en skala med følgende alternativ:

- Very large
- Large
- Neutral
- Small
- Very Small

- Don't know

Alternativer:

- Documentation for the BI-tools used
- Colleagues
- Forums for the BI-tools used
- Customer support from the BI-tools used
- Examples from previous projects you worked on
- Non-tool specific forums for BI developers

Question 18: Have you worked with Open Source stand-alone BI tools (not part of a complete platform)?

- Yes
- No

Dersom respondenten her svarer "yes", blir han ført til Del 4. Dersom respondenten svarer "No" er undersøkelsen ferdig, og han føres til "Avslutning".

A.4 Del 4 - Open Source BI tools

Question 19: When was the last time you worked with Open Source BI tool(s)?

- 1990-1995
- 1996-2004
- 2005-2008
- 2009-2010
- 2011

Question 20: Which tool(s) have you used?

Her kunne respondenten fylle inn svaret i en tekstboks.

Question 21: To what degree do you agree with the following statements?

If you mean the opposite of what the statements says, choose Strongly Disagree.

Her skulle respondenten vurdere hvor enig han var med utsagnene, på en skala med følgende alternativ:

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Alternativ

- Open Source BI is not developed and tested to the same high standards as conventional BI
- Open Source BI products are incomplete
- The open source business model is not viable in the long term
- Open Source BI-tools are less user friendly than conventional BI-tools
- Tool-specific forums are more active for conventional BI-tools than for the case of Open Source BI-tools
- For Small- and middle-sized companies, Open Source BI might be a good choice
- For large companies, Open Source BI might be a good choice

Question 22: Do you think Open Source BI platforms will give commercial BI platforms real competition in the future?

- Yes
- No
- Don't know

A.5 Del 5 - Open Source BI Platforms

For the questions on this page, consider the last Open Source BI project you worked on / is working on.

Question 23: When was the last time you worked on an Open Source BI project?

- 1990-1995

- 1996-2004
- 2005-2008
- 2009-2010
- 2011

Question 24: What platform did you use?

- Pentaho
- Jaspersoft
- SpagoBI
- Actuate BIRT

Question 25: Did you use an enterprise edition of the Platform?

- Yes
- No

Question 26: Why was this platform chosen? (Choose up to three reasons)

- Ability to support large numbers of concurrent users
- Availability of local support
- Bundled with another product
- Chosen vendor did a better sales job
- Completed "proof of concept" faster or better than others
- Corporate standard
- Ease of use for application builders
- Ease of use for end-users
- Fast performance
- Functionality/ product features
- Integrates with other products already in use

- Large data handling capacity
- Low price
- Product reputation
- Range of server platforms supported
- Vendor relationship/ reputation
- Web architecture
- Other (Specify)

Question 27: What was/is your customers primary industry?

- Government/Public sector
- Financial services
- IT and Technology
- Manufacturing
- Professional services
- Consumer goods
- Energy and natural resources
- Telecoms
- Construction and real estate
- Transportation, travel and/ tourism
- Entertainment, media and publishing
- Logistics and distribution
- Education
- Other

Question 28: Approximately how many of your customers employees will regularly be using the BI solution?

- 1-5

- 6-10
- 11-25
- 26-50
- 50-100
- 100-200
- 200-500
- 500-2000
- 2000-5000
- 5000 < Don't know

Question 29: What percent of those BI users would you consider to be power users (i.e. have the ability to drill down, analyze data, and create reports) - as opposed to casual users (periodic viewing of static reports)?

- 0%
- 5%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%
- Don't know

Question 30: What percent of employees with BI access will be using the solution in a self-service, non IT-assisted capacity?

- 0%
- 5%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%
- Don't know

Question 31: What was your customers main motivations for investing in a BI implementation? (Skip if you don't know)

Her skulle respondenten krysse av på en skala med følgende alternativer:

- Very important
- Important
- Neutral
- Unimportant

Alternativer:

- To share information with customers, suppliers and partners
- To move users toward a self service model of information delivery
- To decrease business costs and improve operational efficiency

- To speed up and improve the organization's decision making ability
- To better align with and track against corporate strategy
- To better manage a specific operational process in a more timely fashion
- To respond to user needs for data on a timely basis

Question 32: In which of the following BI activities does your customer engage? (Select all that apply)

- Dashboard visualization of key metrics
- Regular / periodic report generation and delivery
- Ad hoc query and reporting
- Scoreboards and performance management
- Predictive analytics
- Real-time analytics
- Drill-down analytics
- Don't know

Question 33: Which of the following have you been working with on this project, and what licences did the tools have? (Only check the ones you have worked with)

- Data integration
- Meta data management
- Data warehouse
- Development and delivery of reports
- Development and delivery of BI analytic applications
- Development and delivery of dashboards
- Data modeling development
- Cube building

Question 34: To what extent have you used the following sources, when in need of development support?

Her skulle respondenten vurdere hver kilde på en skala med følgende alternativ:

- Very large
- Large
- Neutral
- Small
- Very Small
- Don't know

Alternativer:

- Documentation for the BI-tools used
- Colleagues
- Forums for the BI-tools used
- Customer support from the BI-tools used
- Examples from previous projects you worked on
- Non-tool specific forums for BI developers

Question 35: Please arrange the usefulness of the sources in the previous question.

Her skulle respondenten vurdere nyttheten til hver kilde på en skala med følgende alternativ:

- Very large
- Large
- Neutral
- Small
- Very Small
- Don't know

Alternativer:

- Documentation for the BI-tools used
- Colleagues
- Forums for the BI-tools used
- Customer support from the BI-tools used
- Examples from previous projects you worked on
- Non-tool specific forums for BI developers

Question 36: To what degree do you agree with the following statements? If you mean the opposite of what the statements says, choose Strongly Disagree.

Her skulle respondentene vurdere hvor enig han var med utsagnene, på en skala med følgende alternativ:

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Alternativ

- Open Source BI is not developed and tested to the same high standards as conventional BI
- Open Source BI products are incomplete
- The open source business model is not viable in the long term
- Open Source BI-tools are less user friendly than conventional BI-tools
- Tool-specific forums are more active for conventional BI-tools than for the case of Open Source BI-tools
- For Small- and middle-sized companies, Open Source BI might be a good choice

- For large companies, Open Source BI might be a good choice

Question 37: Do you think Open Source BI platforms will give commercial BI platforms real competition in the future?

- Yes
- No
- Don't know

Respondenten blir tatt videre til siste side av undersøkelsen: "Avslutning".

A.6 Avslutning

Thank you for your participation, your answers will be of great value!

You may also distribute the survey to others with competence in this area. Link to the survey: http://www.kwiksurveys.com/online-survey.php?surveyID=IJDNLL_c602d8f

If you would like to receive the finished Master Thesis in June 2011, please send me a request on hellesof@gmail.com.

Bibliografi

- [1] Aberdeen; *Open Source Business Intelligence*. Aberdeen BeyeNETWORK, 2010. <<http://www.aberdeen.com/Aberdeen-Library/6445/RA-open-source-business-intelligence.aspx>>, aksessert februar 2011.
- [2] Acando Norge hjemmeside: <http://acando.com/About-Acando>.
- [3] All, A.; *SMB Interest in BI Is Growing*. IT BusinessEdge, 2007.
- [4] Baird, S, A.; *The Heterogeneous World of Proprietary and Open-Source Software*, The University of Hong Kong, 2008.
- [5] BARC; *The BI Survey 9 - The customer verdict*. BARC; Business Application Research Center, 2010.
- [6] Bickman, L; Rog, D. J.; *Handbook of applied social research methods* , Sage Publications, Thousand Oaks, 1998.
- [7] Bitterer, A.; *Who's Who in Open-Source Business Intelligence*, Gartner, 2008.
- [8] Bouman, R.; Dongen, J, V.; *Pentaho Solutions - Business Intelligence and Data Warehousing with Pentaho and MySQL*. Wiley Publishing, 2009.
- [9] Businessday; <<http://www.businessday.com.au/small-business/blogs/work-in-progress/sizing-up-small-employers/20110311-1bpw4.html>>, aksessert januar 2011.
- [10] Chen, R, S.; Sun, C, M.; Helms, M, M.; Jih, W, J.; *Role Negotiation and Interaction: An Exploratory Case Study of the Impact of Management Consultants on ERP System Implementation in SMEs in Taiwan, Information Systems Management*, 2008.

- [11] Clarke, A.; Dawson, R.; *Evaluation Research: An introduction to Principles, Methods and Practice*, Sage Publications, London, 1999.
- [12] Cooper, R. B.; Zmud, R.W.; *Information Technology Implementation Research: A Technological Diffusion Approach*, Management Science, 1990.
- [13] Curtis, B.; Krasner, H.; Iscoe, N.; "A field study of the software design process for large systems". Communications of the ACM, 1988.
- [14] Daffara, C.; González-Barahona, J. M.; Humenberger, E.; Koch, W.; Lang, B.; Laurie, B.; *Free Software / Open Source: Information Society Opportunities for Europe?* Information Society Directorate General of the European Commission, 2000.
- [15] Davenport, T. H.; Prusak, L.: *Working Knowledge: How organizations manage what they know*. Harvard Business School Press, 2000.
- [16] Eckerson, W. W.; *Performance Dashboards - Measuring, Monitoring and Managing Your Business*. Wiley, 2006.
- [17] Electrosmart; *BI For Small and Medium Businesses*. The Business Intelligence Guide, 2006. <http://thebusinessintelligenceguide.com/bi_strategy/BI_For_SMBs.php> aksesert 12 februar, 2011.
- [18] EU; SME definitions, 2003. <http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/index_en.htm>, aksesert januar 2011.
- [19] Fiedler, G.; Raak, T.; Thalheim, B.; *Database Collaboration Instead of Integration*. Computer Science and Applied Mathematics Institute, Kiel University, 2005.
- [20] Fisher, D. M.; Kiang, M. Y.; Fisher, S. A.; Chi, R. T.; *Evaluating mid-level ERP software*, Journal of Computer Information Systems, 2004.
- [21] Fitzgerald, B.; *The transformation of open source software*, MIS Quarterly, 2006.
- [22] Flowers, B.; *Business Intelligence - An Analytical Opportunity?* Requirements Network, 2009.
- [23] Fogel, K.; *Producing Open Source Software. How to Run a Successful Free Software Project*, 2007.

- [24] Forskningsrådet; *Guide til den nye SMB-definisjonen.* <http://www.forskningsradet.no/no/Nyheter/Guide_til_den_nye_SMB_definisjonen/1236685416549>, akessert januar 2011.
- [25] Fossum, F.; Hansen, M, A.; Hemmen, P.; Karud, C, H.; Torgersen, M, N.; *Business Intelligence on an Open Source-platform.* TDT4290-Kundestyrt Prosjekt, NTNU, 2009.
- [26] Franke, N.; Hippel, V, E.; *Satisfying heterogeneous user needs via innovation toolkits: The Case of Apache Security Software.* MIT Sloan School of Management, 2002.
- [27] Gartner; *Overcome the four IT inhibitors to BI success in midsize companies.* IBM, 2010.
- [28] Golfarelli, M.; *Open Source BI Platforms: A Functional and Architectural Comparison.* DEIS, University of Bologna, 2002.
- [29] Grimes, S.; *MySQL V5 – Ready for Prime Time Business Intelligence.* Alta Plana Corporation, 2006.
- [30] Hertel, G.; Niedner, S.; Herrmann, S.; *Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel.* Research Policy in press, 2003.
- [31] Hippel, V, E.; Krogh, V, G.; *The private-collective innovation model in open source software development.* Organization Science, 2003.
- [32] IT-Novum. *Open Source Business Intelligence - A comparison of the Open Source BI Solutions Jaspersoft, Palo and Pentaho.* IT Novum GmbH White Paper, 2009.
- [33] Kanywani, M.; *One-on-One with Pentaho,* 2008.
- [34] Kernochan, W.; *What Is Business Intelligence and What Does It Mean for the SMB?* Focus, 2010.
- [35] Klein, H, K.; Myers, M, M.; *A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems.* MIS Quarterly, 1999.
- [36] Krogh, G. V.; Spaeth, S.; Lakhani, K. R. *Community, joining, and specialization in open source software innovation: a case study.* Institute of Management, University of St. Gallen, 2003.

- [37] Krogh, G, V.; Hippel, E, V.; *Special issue on open source software development*, Research Policy, 2003.
- [38] Kwiksurvey; <<http://www.kwiksurveys.com>>, aksessert februar 2011.
- [39] Lakhani, K, R.; Wolf, R, G.; *Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects*. MIT Sloan School of Management, 2005.
- [40] Larson, B.; *Delivering Business Intelligence with Microsoft. SQL Server 2005. Essential skills for database professionals*. McGraw-Hill, 2006.
- [41] Laukkanen, S.; Sarpola, S.; Hallikainen, P.; *ERP System Adoption – Does the size Matter?*, *Proceedings of the 38th. Hawaii International Conference on System Sciences*, 2005.
- [42] McDonald, K.; Wilmsmeier, A.; Dixon, D, C.; Immon, W, H.; *Mastering the SAP Business Information Warehouse: Leveraging the Business Intelligence Capabilities of SAP NetWeaver*. Wiley, 2002.
- [43] Mockus, R, T.; Herbsleb, F, J.; *Two case studies of open source software development: Apache and Mozilla*, ACM ToSEM, 2002.
- [44] Monteiro, E.; *Integrating Health Information Systems: A Critical Appraisal*. Institutt for Datateknikk og Informasjonsvitenskap, Norges teknisk- og naturvitenskapelige universitet (NTNU), 2003.
- [45] Monteiro, E.; Forelesning i faget IT3604- IKT og Organisasjon. *Tema: Open Source*, 21 Okt. 2009.
- [46] Monteiro, E.; Jæger, J, F.; *Realizing organizational benefits with ICT in healthcare: the challenge of integration*. Institutt for Datateknikk og Informasjonsvitenskap, Norges teknisk- og naturvitenskapelige universitet (NTNU), 2005-
- [47] Moss, L, T.; Atre, S.; *Business Intelligence Roadmap*. Addison, Wesley, 2003.
- [48] Mundy, J.; Thornthwaite, W.; Kimball, R.; *The Microsoft Data Warehouse Toolkit. With SQL Server 2005 and the Microsoft Business Intelligence Toolset*. Wiley Publishing Inc, 2006.
- [49] Myers, M, D.; *Investigating information systems with ethnographic research*. University of Auckland, 1999.

- [50] Myers, M, D.; Newman, M.; *The qualitative interview in IS research. Examining the craft*. University of Auckland, 2006.
- [51] Nelson, R, R.; *IT Project Management: Infamous Failures, Classic Mistakes, and Best Practices*. University of Minnesota, 2007.
- [52] Oates, B, J.; *Researching information systems and computing*. SAGE publications, 2006.
- [53] Pentaho; *Pentaho Open Source Business Intelligence Platform Technical White Paper*. Pentaho Corporation, 2006.
- [54] Pentaho webside; *Pentaho amps up leading energy firms*, Pentaho Website, 2010. <http://www.pentaho.com/news/releases/20100819_pentaho_amps_up_leading_energy_firms.php>, aksessert juni 2010.
- [55] Pentaho webside; *Pentaho plays lead role in TV2s sumo show*, Pentaho website, 2010. <http://www.pentaho.com/news/releases/20100624_pentaho_plays_lead_role_in_tv2s_sumo_show.php>, aksessert juni 2010.
- [56] Ramakrishnan, R.; Gehrke, W, J.; *Database Management Systems*, McGraw-Hill, 2003.
- [57] Regjeringen; *Små og mellomstore bedrifter* <http://www.regjeringen.no/nb/dep/nhd/tema/forenkling_for_naringslivet/sma-og-mellomstore-bedrifter.html?id=614069>, aksessert januar 2011.
- [58] Rødsjø, B.; Benjaminsen, H, S, T.; *Rapport for Sommerjobb 2010*, Acando, 2010.
- [59] Rødås, K.; Stoveland, J, F.; *Business Intelligence Competency Centers*, Capgemini Norge, 2010.
- [60] Socialresearchmethods; <<http://www.socialresearchmethods.net>>, akessert november 2010.
- [61] Sourceforge; <<http://sourceforge.net/about>>, aksessert mars 2011.
- [62] Spørreundersøkelse; <http://www.kwiksurveys.com/online-survey.php?surveyID=IJDNLL_c602d8fd>, Kwisurvey, 2011.

- [63] Swanson, E.; Beath, A.; “*The use of case study data in software management research*”. Journal of Systems and Software, 1988.
- [64] Thong, J, Y, L.; Chee-Sing, Y.; Raman, K, S.; *Top Management Support, External Expertise and Information Systems Implementation in Small Businesses*, Information Systems Research, 1996.
- [65] Tucci, T.; *Business intelligence now a given for SMBs*. ComputerWeekly, 2007.
- [66] Thomsen, C.; Pedersen, T, B.; *A Survey of Open Source Tools for Business Intelligence*, DB Tech Reports, 2008.
- [67] Utley, C.; *Designing the Star Schema Database*, 1999
- [68] Vision Solutions; *Get the Facts: Real-Time, On-Demand Information for Business Intelligence and Data Integration*. Vision Solutions Inc, 2009.
- [69] Walsham, G.; *Knowledge Management: The benefits and limitations of computer systems*, 2001.
- [70] Ward, R.; *Data Integration*. Strategic Procurement Services, 2004.
- [71] Watsom, H, J.; Wixom, B, J.; *IEEE: The Current State of Business Intelligence*. IEEE Computer Society, 2007.
- [72] Wheeler, D, A.; *Why Open Source Software / Free Software (OSS/FS, FLOSS, or FOSS)? Look at the Numbers!*, 2007.
- [73] Whitehorn, M.; Zare, R.; Pasumansky M.; *Fast track to MDX*, Microsoft Corporation, 2002.
- [74] Wikipedia; *Empirical Research*. <http://en.wikipedia.org/wiki/Empirical_research>, aksessert oktober 2010.
- [75] Wikipedia; *Small and Medium businesses*. <http://en.wikipedia.org/wiki/Small_and_medium_businesses>, aksessert januar 2011.
- [76] Wise, L.; *Analyst Perspectives - A General Look at Business Intelligence for Mid-market*. Beye NETWORK, Global coverage of the business intelligence ecosystems, 2010.
- [77] Ye, Y.; Kishida, K.; *Toward an understanding of the motivation of Open Source Software Developers*, IEEE Computer Society Press, 2003