

EXAMENSARBETE

Data Warehouse som beslutsstöd

Problem för slutanvändare

LISELOTTE HOLMBOM
TERESE LASSINANTTI

Samhällsvetenskapliga och ekonomiska utbildningar

SYSTEMVETENSKAPLIGA PROGRAMMET • C-NIVÅ

Institutionen för Industriell ekonomi och samhällsvetenskap
Avdelningen för Systemvetenskap • Data och systemvetenskap

Sammanfattning

Syftet med vårt arbete har varit, att utröna vilka problem som kan uppstå för slutanvändare av datalager. Vi har även frågat oss vad dessa problem kan bero på. För att uppnå detta, har vi först gjort en litteraturstudie och sedan genomfört intervjuer med slutanvändare och IT-tekniker på fyra olika företag runtom i Sverige. Undersökningen har visat, att det finns en hel del problem förknippade med användning av datalager och att man också kan se några framträdande sådana, såsom exempelvis långa svarstider vid sökningar. De problem som visat sig är dock av sådan karaktär, att de vida övervägs av de fördelar ett datalager ger åt företag. Vi tycker ändå det är viktigt att notera dem.

Abstract

The aim of our work has been to show what problems can arise for end-users of data warehouses, and what causes these problems. To accomplish this task, we first undertook literature studies, and then performed interviews with end-users and IT-technicians from four different companies around Sweden. Our investigation has shown, that there are many problems associated with the use of data warehouses, some of which are more dominant than others - for example, long response times for queries. The problems that have shown are though small enough to be overshadowed by the huge benefits that a data warehouse gives to the company. Nevertheless, they deserve our attention.

Förord

Denna rapport är resultatet av ett examensarbete i systemvetenskap omfattande 10 poäng, som ett sista moment i utbildningsprogrammet för Data- och Systemvetenskap vid Luleå Tekniska Universitet. C-uppsatsen ger, tillsammans med övriga kurser, en kandidatexamen om 120 poäng inom ämnet.

Med anledning av arbetets genomförande, vill vi speciellt tacka följande företag och personer: vår handledare Sören Samuelsson (för nyttiga tips och råd), läraren Ingela Johansson (för att hon upplåtit tid och arbetsrum åt oss när vi behövt i samband med intervjuer), Patrik Hägglund ("teknisk assistans", på obekvämt arb.tid!), vänner som testläst och kommenterat, samt Stadium, Frontec, Findus, Vattenfall (för att vi fått göra undersökningen hos dem med så positivt bemötande), alla de företag som tagit sig tid att svara på vår enkät och hela seminariegruppen (för mycket konstruktiv respons).

Luleå 03-06-24

.....
Liselotte Holmbom

.....
Terese Lassinantti

Innehållsförteckning

1	BAKGRUND	7
1.1	INLEDNING	7
1.2	SYFTE OCH FORSKNINGSFRÅGA	10
1.3	AVGRÄNSNINGAR	10
2	TEORI	12
2.1	VAD ÄR DATA MINING?	15
2.2	VAD ÄR OLAP?	16
2.3	SLUTANVÄNDARE OCH DATALAGER	18
2.3.1	<i>Medverkansbrister</i>	18
2.3.2	<i>Bristfälligt användarstöd</i>	20
2.3.3	<i>Datautvinningsproblem</i>	22
2.3.4	<i>Datapresentationsproblem</i>	23
3	METOD	24
3.1	LITTERATURSTUDIE	24
3.2	UNDERSÖKNINGSMETOD	24
3.2.1	<i>Datainsamling</i>	25
3.3	ANALYSMETOD	25
3.4	VALIDITET OCH RELIABILITET	26
3.5	SAMMANFATTNING AV VÅRT ARBETE:	28
4	EMPIRI	29
4.1	ENKÄTSAMMANSTÄLLNING	29
4.1.1	<i>Företags- och respondentbeskrivning – intervjuobjekt</i>	30
4.2	SAMMANSTÄLLNING AV INTERVJUSVAREN	33
5	ANALYS	37
5.1	ENKÄTSVAR	37
5.2	INTERVJUSVAR - ÖVERSIKT	38
5.3	INTERVJUSVAR - DETALJANALYS	39
6	DISKUSSION OCH SLUTSATSER	46
6.1	LITTERATUR	46
6.2	FALLSTUDIE	46
6.3	UNDERSÖKNINGSSPECIFIKA SLUTSATSER	47
6.4	ALLMÄNNA SLUTSATSER	47
6.5	SUMMERING AV EGNA REFLEKTIONER	48
6.6	FÖRSLAG TILL FORTSATT FORSKNING	50
7	LITTERATURFÖRTECKNING	51
	BILAGA A - ENKÄTFRÅGOR	52
	BILAGA B - INTERVJUFRÅGOR	53

Figur- och tabellförteckning:

<i>Figur 1: Datalager och slutanvändare</i>	9
<i>Figur 2: Datalager</i>	13
<i>Figur 3: OLAP-kub</i>	17
<i>Figur 4: Ökande krav/komplexitet av analys.</i>	21
<i>Figur 5: Metodmodell.</i>	28
<i>Tabell 1: Enkätansammanställning</i>	29
<i>Tabell 2: Medverkansbrister</i>	33
<i>Tabell 3: Bristfälligt användarstöd</i>	34
<i>Tabell 4: Datautvinningsproblem</i>	35
<i>Tabell 5: Datapresentationsproblem</i>	36
<i>Tabell 6: Användarproblem från enkätsvar</i>	37
<i>Tabell 7: Intervjusvar ja/nej</i>	38

1 Bakgrund

1.1 Inledning

Det finns stora förtjänster med datalager för de företag som tillämpar det. Connolly & Begg (1999) tar upp följande **fördelar med ett datalager**:

- En lönande investering (det är rätt så dyrt att bygga ett datalager, men de pengar som går åt till investeringen vinnns tillbaka många gånger om)
- Marknadsvinster
- Ökad produktivitet hos företagets beslutsfattare

Det gäller för företag att utnyttja de möjligheter som ett datalager ger. Ett datalager ger företag en möjlighet att få större kunskap om sin verksamhet, bland annat genom att man kan se samband, fakta om kunder och leverantörer, samt se olika trender. Dessutom ger datalager marknadsfördelar gentemot konkurrenter, produktiviteten ökar för beslutsfattare inom företaget (d v s de kan utföra mer på kortare tid), samt att kostnaden för datalagret betalar sig i form av informations- och tidsvinster (Söderström, 1997).

Enligt Söderström (1997) kan ett **datalager bland annat lösa följande problem**:

- Data är svårt att förstå, då man inte använder sig av några andra termer än de som finns i verksamheten.
- Data är felaktigt uppbyggt och kan därför inte precist belysa aktuella verksamhetsproblem.
- Det går inte att flexibelt söka och sammanställa data på det sätt verksamheten kräver.
- Ingen jämn och känd kvalitet finns på data.
- Det går inte att jämföra data från olika källor.
- Man kan inte göra jämförelser mellan olika tidsperioder.
- All den data man behöver som beslutsunderlag finns ej tillgänglig för den som ska fatta beslut.
- Beslutsfattaren kan inte själv utan hjälp begära fram och sammanställa de data som behövs i en given situation.

När man läser om datalager framkommer, att det inte bara finns fördelar med dessa, utan de är även förknippade med en rad problem. De undersökningar som tidigare gjorts inom området pekar nästan enbart på de enorma förtjänster datalager ger för företaget. Hortell & Soini (1999) hade som uppdrag att undersöka fördelar, men noterade ändå några mindre bra saker (t ex den väldigt korta introduktion användare ges i sitt verktyg). Vi väljer därför att angripa de **problem** som finns med användandet av datalager.

En datalagerlösning beställs inte som ur en katalog, utan tas fram i utvecklingsprojekt, antingen av företaget själva eller av inhyrda IT-konsulter. Datalagerprojekt kan pågå i flera år, det är alltså inget man tar fram över en natt. Det är viktigt att man kommer in rätt från början, för att slutresultatet ska bli bra.

Varför misslyckas datalagerprojekt?

Data Administration Newsletter (2001) av Robert S. Seiner tar upp elva vanliga anledningar till att datalagerprojekt betraktas som misslyckade. Av dessa finner vi följande särskilt intressanta: 1) Missnöjda användare, 2) Funktioner och kapaciteter blev aldrig implementerade, 3) Oacceptabel "performance" och 4) För komplicerat för användarna.

1) Om användarna är missnöjda ska projektet anses misslyckat. Missnöjdhet är ofta ett resultat av orimliga förväntningar - något fick användarna att förvänta sig mer än vad de fick, ofta beroende på ett "glapp" i kommunikationen mellan IT-avdelningen och användaren.

2) I början av projektet satte man upp särskilda mål för funktion och kapacitet. Dessa kan ha inkluderat datakvalitet, användarutbildning, sättet att leverera data, eller fördefinierade frågor. Om sådana funktioner aldrig realiserades, eller blev uppskjutna till senare tillfällen, kan det betraktas som ett misslyckande.

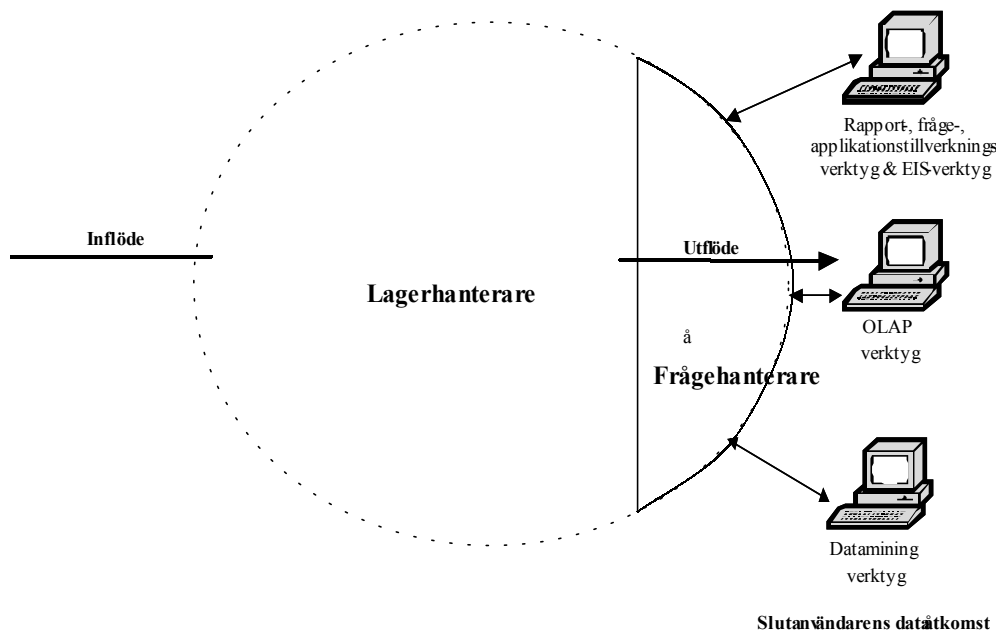
3) Datalagerprojekt avblåses ofta på grund av oacceptabel performance. Performance syftar både på svarstid för frågor och på tid för inladdning av data till datalagret. Användaren måste förstå vad man kan förvänta sig, men samtidigt finns det gränser för vad man skall acceptera! Ju större ett datalager blir, desto längre brukar svarstiderna bli.

4) En del verktyg är för svåränvända för beslutsfattaren. IT-avdelningen väljer ofta verktyg som de själva är förtrogna med, men detta gäller inte för slutanvändarna! Om verktyget är för invecklat, kommer slutanvändaren att försöka undvika att använda det. Detta stjälper själva syftet med ett datalager, att slutanvändaren ska uppmuntras att använda frågeverktyget självständigt.

Idag fyller företag sina databaser med data som de tror sig ha nytta av. Problemet är svårigheterna med att analysera och använda all den data man samlat in (Söderström, 1997). Till exempel tycks det så, att när slutanvändarna av datalagret får sina fråge- och rapportverktyg, ökar deras behov av hjälp från teknisk personal snarare än minskar (Connolly & Begg, 1999). Med slutanvändare menas alltså de personer som tar fram data, bearbetar den och fattar de strategiska besluten i ett företag. Det är meningen att verktygen ska **uppmuntra användarna att själva ställa komplexa frågor**, men de är ofta utformade på ett sådant sätt att slutanvändaren inte klarar att göra sökningen själv. Dessutom är många slutanvändare inte skolade i att använda verktygen - idag är det vanligt att företagen bara ger en kort utbildning i verktygen eller i gränssnittet, sedan får användaren sitta och lära sig själv. Detta är många gånger otillräckligt och kan

exempelvis leda till att data blir svårtolkad/svår att förstå för användaren "Fenomenet" framkommer i de rapporter vi läst.

Användare blir dock alltmer medvetna om hur mycket information man bör kunna utvinna ur ett datalager, men förmår alltså ibland inte själva ta fram den. Detta problem kan delvis avhjälpas genom att köpa in verktyg som är lättare att använda, verktyg som har större kapacitet, samt att ge användarna mer utbildning i verktygen (Connolly & Begg, 1999).



Figur 1: Datalager och slutanvändare

Källa: omarbetad bild från Connolly & Begg (1999)

Vi vill undersöka situationen hos svenska företag idag och se om de har problem ur slutanvändarperspektiv vid tillämpningen av datalagret. Söderström (1997) menar, att en del problem inte beror på verktygen utan har mer att göra med att användarna inte fått tillräcklig utbildning, inte fått vara med vid framtagningen av datalagerlösningen, att verktygen inte är anpassade efter den specifika användaren eller till verksamheten i stort, eller att användargränssnittet är dåligt utformat. Med denna rapport vänder vi oss i första hand till tillverkare av datalagerlösningar samt slutanvändare, för att göra dem uppmärksamma på att det finns sådana här problem. Det tycks inte som att många ställt sig denna fråga tidigare, men det är nödvändigt för att användarna ska kunna tillämpa datalagret optimalt. Rapporten riktar sig även till andra intresserade, såsom studenter som skall jobba med denna problematik senare.

Här är några begrepp som vi kommer att använda förklarade:

Analysverktyg – En produkt som ger slutanvändarna direkt åtkomst och möjligheter till manipulation av data inom systemet. Detta skall ske utan någon som helst inblandning av programmeringsarkivet.

Källsystem – Operativa system för daglig verksamhet från vilka data hämtas till datalagret (Söderström (1997, s.198).

Metadata – Data om data (Söderström 1997, s.198).

Snöflingestruktur – Påminner mycket om stjärnstruktur (se nedan) men här är det flera olika nivåer av tabeller som pekar mot faktatabellen (Söderström 1997, s. 198).

Stjärnstruktur – En struktur som liknar en stjärna, med en stor tabell (faktatabell) med summerbara uppgifter i mitten och flera så kallade dimensionstabeller runt om som medger sammanställning av data efter olika indelningsgrupper (Söderström 1997, s. 198).

Strategiska beslut – Beslut som fattas på ledningsnivå i ett företag eller i en organisation. Beslut som rör den strategiska styrningen och kommer att ha stor betydelse för hur framtiden händelser i företaget. Är ofta beslut som tas på långsikt, eftersom konsekvenserna av dessa beslut ofta berör många människor, både människor som finns i företaget men eventuellt även människor utanför företaget.

1.2 Syfte och forskningsfråga

Syfte: Att ge en samlad bild av förekommande svårigheter för slutanvändare av datalager.

Forskningsfråga: Vilka problem kan uppstå för slutanvändare när de tillämpar datalagret och vad beror dessa problem på?

1.3 Avgränsningar

- Vi valde att intervjua **vinstdrivande företag**, som **använder datalagret till strategiskt beslutsfattande**. Detta för att de är många och använder sina datalager på ett likartat sätt - de samlar data som hör till vardagen, t ex inköp, och som därför berör många.
- Vi undersökte endast de **problem kopplade till informationsutvinningen** ur datalagret och utelämnade problem av annan karaktär.

- Vi har medvetet intervjuat företag som är i **olika branscher**, för att se om man har samma typer av problem fast man bedriver helt olika verksamhet. Dessa företag kan ha **olika användarverktyg** till sina datalager – vi har inte avgränsat bort någon verktygskategori, då vi inte vill påvisa brister hos ett visst verktyg.

2 Teori

Ända sedan 1970-talet har företag fokuserat sina investeringar i nya datasystem som automatiserar affärsprocesser. Företag och organisationer vill finna sätt att använda operativ data som beslutsstöd, för att vinna marknadsandelar (Connolly & Begg, 1999).

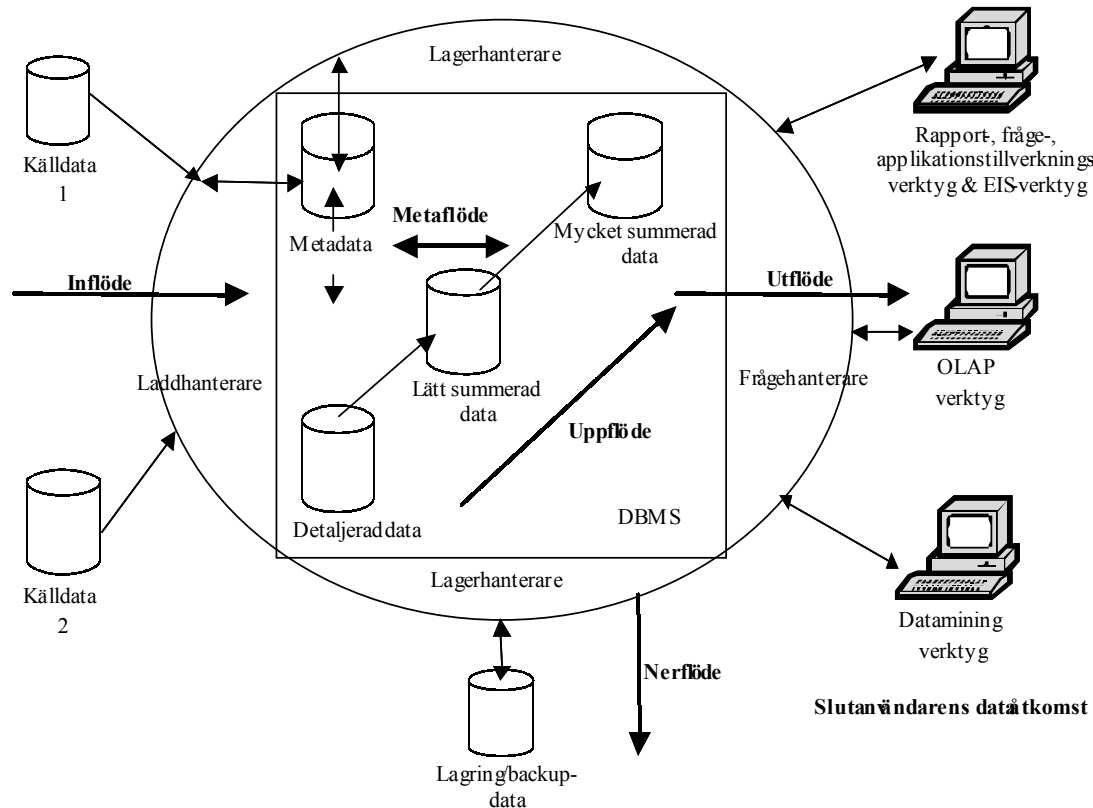
För att ett företag skall kunna fatta bra och snabba beslut krävs ett bra beslutsunderlag. Ett datalager är bra för att ge just detta – tillhandahålla den information beslutsfattare inom företaget behöver. Under de senare åren har det blivit allt vanligare med datalager hos större företag, som en hjälp för att ställa prognoser och fatta beslut. Hjälpen består av sammanställd data över exempelvis försäljningssiffror sedan flera år tillbaka (kundprofiler, m m). Detta gör att de strategiska besluten blir bättre, eftersom de grundas på uppgifter från hur verksamheten gått tidigare (Söderström, 1997).

Det finns många olika **definitioner av vad ett datalager är**. Här är en:

”Ett datalager är en logiskt sammanhållen datamängd, som är avsett för analys som speglar flera tidsperioder genom att data regelbundet hämtas från andra register” (Söderström, 1997, s.197).

Det övergripande målet med ett datalager enligt Connolly & Begg (1999) är att integrera företagsdata till en samlad källa, från vilken användare lätt kan köra sökningar, skapa rapporter och utföra analyser. Ett datalager är en beslutsstödjande miljö, som samlar data lagrade på olika källor, organiserar den och gör den tillgänglig för beslutsfattare i organisationen. Man kan kort säga, att ett datalager är datahanterings- och data-analysteknologi! För att göra det lättare för beslutsfattare, **paketeras och anpassas informationen efter de behov** beslutsfattarna har - detta kan man göra med en så kallad **datamart**. En datamart är en lokal databas, till vilken data från datalagret kopieras. Varje avdelning kan till exempel ha en egen datamart, där endast för just den avdelningen relevant data finns (Söderström, 1997).

En datalagermiljö utgörs av ett antal delar:



Figur 2: Datalager

Källa: omarbetad bild från Connolly & Begg (1999)

Förklaring till bilden med de olika flödena:

Inflöde– processerna som har med överföring, tvätt och inladdning av data från källsystemen att göra.

Uppflöde – processerna som har att göra med tillförande av värde till data i datalagret, genom summering, hopslagning och distribution av den.

Nerflöde– processerna som har med arkivering, backuptillverkning och återställande efter skador i datalagret att göra.

Utföde – processer associerade med att göra data tillgängligt för slutanvändare.

Metaflöde – processerna som har med hanteringen av metadata att göra.

Datalagerhållning är en process för att samla in data från olika delar i företaget, se till att data har rätt format och sedan ladda in den i databasen. Data i denna databas kan sedan analyseras och granskas på olika sätt med hjälp av verktyg.

Att bygga ett fullt integrerat datalager är ett komplext arbete, eftersom det inte finns någon försäljare som har ett ”end-to-end”-parti av verktyg. Detta betyder att ett datalager är byggt av flera produkter, vanligen från olika leverantörer. Att få dessa produkter att fungera bra ihop och få dem fullt integrerade är en stor utmaning. Detta fungerar av förklarliga skäl inte alltid utan problem.

Connolly & Begg (1999) tar upp följande **problem med datalager**:

- **Underskattning av resurser för datainladdning** – Det är lätt att underskatta den tid som går åt för att utvinna, tvätta och ladda data till datalagret. Denna laddningsprocess tar upp cirka åttio procent av den totala tiden vid utveckling av ett datalager.
- **Gömda problem med källsystemen** – Problem som har med inmatningen i källsystemen i datalagret att gör kan vara svåra att upptäcka och märks oftast först efter några år.
- **Nödvändig data finns ej tillgänglig** – Datalagerprojekt har ofta ett behov av data som inte tillvaratas av de existerade källsystemen.
- **Ökat slutanvändarberoende** – Efter att slutanvändarna av datalagret får sina fråge- och rapporteringsverktyg, ökar deras behov av support från IT-personal, istället för att - som meningen är - minska.
- **Datahomogenitet** – Stora mängder data kan fresta datadesignern att se (och eftersträva) likheter istället för skillnader, vilket minskar värdet av datat.
- **Stort beroende av resurser** - Ett datalager kan ta upp stort diskutrymme. Många datalager som används för beslutsstöd är designade efter stjärnstruktur eller snöflingstruktur. Dessa strukturer gör att man måste skapa stora fakta tabeller. Om, till exempel, den faktiska datan är i flera dimensioner, eller är kombinationer av tabeller, kan index till den faktiska datan ta upp mer utrymme än rådatan!
- **Dataäggande** – Data som bara en viss avdelning haft tillgång till, blir nu synlig för fler – kan vara svårt att acceptera.
- **Långvariga projekt** – Att bygga ett datalager kan ta flera år.
- **Komplexitet av integration** – Den viktigaste delen i hanteringen av ett datalager är integrationen av de delar som skall åstadkomma kapaciteten.

De punkter vi fäster speciellt stor vikt vid i vår utredning är de sex sista. Dessa har på olika sätt med datautvinningen att göra – inte sällan visar sig problem inom dessa områden just under datautvinningen. Problem av slaget ”ökat användarberoende” kan

exempelvis visa sig som svårigheter att använda analysverktyget eller förstå den utvunna datan - något som i sin tur kan vara en följd av att man ej inkluderat användarna i datalagerprojektet (se kap. 2.3.1).

För ett företag är det inte nog att bara ha stora mängder data lagrat i datalagret. För att få användbar information krävs mänsklig tolkning eller någon annan typ av bearbetning. För att fatta bra beslut måste man ha lämpliga verktyg för att hitta och utvinna specifik data och sedan omvandla den till verksamhetsinformation och övervaka förändringar. Problemet är ofta inte brist på data, utan svårigheter i hur man ska söka i de stora mängderna data för att hitta användbar information (Söderström, 1997). Detta kan lösas med hjälp av bra analysverktyg - verktygen automatiserar sökningsprocessen och presenterar verksamhetsinformation. Det är alltså mycket viktigt att ha rätt analysverktyg (Connolly & Begg, 1999).

Det analysarbete användarna utför med hjälp av verktygen går i stort ut på följande:

Man formulerar en fråga om det man vill veta om, t ex kundstatistik; ”Hur många gånger dagligen ringer kvinnor mellan trettio och trettiofem i fyrkantområdet mobilsamtal där mottagaren betalar?”. Man bestämmer sedan vilka tabeller i datalagret som skall kunna svara på denna fråga och skapar sedan en så kallad kub med dessa tabeller. (Denna del av analysarbetet har många användare dock befriats från, då det är möjligt att förhandsdesigna kuben för varje användare för att underlätta sökningsprocessen – man vet t ex vilken information och vilka tabeller en viss användare ofta behöver). Sedan gör man sökningen, som - beroende på hur komplex frågan är - kan ta olika tid. I kuben med dess dimensioner läggs svarsdata in. Svaret levereras på dataskärmen och kan också skrivas ut som en rapport – antingen av verktyget själv, då presentationen mest blir i form av grafer, eller genom att klippa in resultatet i ett Excel-dokument, där det istället presenteras i form av tabeller med ord (många tycker det är lättare att tolka och göra jämförelser mellan olika slags data och tidsperioder på detta sätt).

Det finns ett par huvudgrupper av sådana här verktyg; **Data Mining** och **OLAP**.

2.1 Vad är Data Mining?

Syftet med Data Mining är att hitta information som är gömd och oväntad. Det är en ganska ny teknik, men som växer snabbt. Data Mining är inget enhetligt område, utan består av ett antal tekniker som är bra på att hitta mönster i underliggande data och i ett datalager med mycket stora datamängder. De kan ”ställa prognoser”(Söderström, 1997).

Data Mining-teknikerna kan bestå av lärande och icke lärande system. I de lärande systemen kan datan ”själv” bygga modeller, medan analytikern i de icke lärande systemen måste ställa upp den modell som systemet sedan använder, samt att användaren måste förstå hur systemet resonerat för att komma fram till resultatet -

”spårbarhet” (Söderström, 1997). Analysverktyget gräver sig ner genom data och det är viktigt att man kan följa hela vägen ner till detaljnivå. Den information man får fram med hjälp av Data Mining är sådant som man normalt inte kan komma åt med vanliga frågor och rapporter. Skillnaden mellan OLAP och Data Mining finns alltså här - OLAP kan bara få fram information om sådant som redan hänt, d v s historiska data (Söderström, 1997).

”En trend är” (dock) ”att klientverktygen mer och mer liknar varandra och kategorierna suddas ut” (Söderström, 1997, s.13).

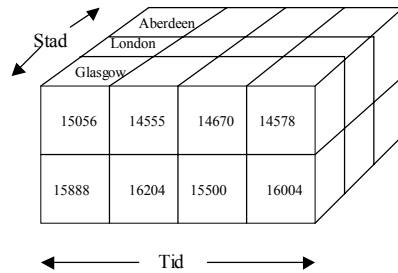
2.2 Vad är OLAP?

E. F. Codd skapade 1993 detta multidimensionella analysverktyg – ”OnLine Analytical Processing” - för sammanställning av data. Det finns i dag många olika produkter, av vilka många kan köras i en vanlig pc och ändå klara att hantera stora datamängder (Söderström, 1997).

Vanliga relationsdatabaser som har rader och kolumner (två-dim.) kan inte ge den multidimensionella syn på data som krävs för att göra utförliga och bra analyser. OLAP-databasservrar använder multidimensionella strukturer för att lagra data och relationer mellan dem. Multidimensionella strukturer är bäst visualiserade som kuber av data och kuber i kuber av data (Söderström, 1997). Varje sida av en kub är en dimension. Multidimensionella databaser bjuder ett kompakt och lättförståeligt sätt att visualisera och manipulera dataelement som har många inbördes relationer (Connolly, 1999). OLAP bygger på att beslutsfattaren ur datalagret bestämmer vilken data man är intresserad av och förberäknar dessa data till så kallade OLAP-kuber (av databasdimensioner), som uppdateras kontinuerligt (se bild med exempeldimensioner!). Med dessa kuber kan användaren skapa dynamiska rapporter direkt på skärmen. Det går sedan snabbt att göra sammanställningar, hitta samband och se mönster. Med OLAP-kuben i figur 3, kan man exempelvis redovisa svaret på följande fråga (om kuben ”lodrätt” visar **typer av fastigheter**, samt **kvartalsresultat (stad för stad och för villa/insatslägenhet)**): **”Hur stora är de årliga intäkterna från all fastighetsförsäljning i samtliga städer?”** (Svaret/resultatet av sökningen är en **summa/produkt** av värdena i kubens dimensioner.)

Med de flesta OLAP-verktyg kan man göra analyserna direkt över Internet, för att slippa sitta vid själva databasen.

OLAP-verktygen är indelade i följande undergrupper: MOLAP (Multi-dimensional OLAP), ROLAP (Relational OLAP), samt MQE (Managed Query Environment). Var och en av dessa har dock, jämte fördelarna, mindre bra egenskaper (Connolly, 1999). MOLAP har t.ex. begränsade funktioner för navigering och dataanalys, på grund av att datan måste designas efter förutbestämda ”användarmönster”. När det gäller ROLAP, så måste man utveckla ett mellanlager som omvandlar den tvådimensionella relationstabellen till en multidimensionell struktur. I MQE kan endast en begränsad mängd data hanteras effektivt (Connolly & Begg, 1999).



Figur 3: OLAP-kub

Källa: omarbetad bild från Connolly & Begg (1999).

Då det tycks vara vanligast - i vart fall i Sverige - att företag använder sig av **OLAP-verktyg** för dataåtkomst, finner vi följande **tolv regler av Codd**, för **urval av OLAP-verktyg**, relevanta:

- 1) **Multidimensionell konceptuell översikt** – OLAP-verktyg skall ge användarna en multidimensionell vy, som motsvarar deras bild av företaget och som är ”intuitivt analytisk” och lätt att använda.
- 2) **Transparens** - med OLAP-teknologi skall den underliggande databasen och arkitekturen vara dold för användarna - detta för att värna användarens produktivitet och tillförlitlighet till kundanpassade applikationer.
- 3) **Åtkomst** – OLAP verktyget ska kunna komma åt nödvändig data från alla olika heterogena företagsdatakällor.
- 4) **Konsistens i ”performance”** – antalet dimensioner, aggregationsnivåer och storleken på databasen ska inte påverka verktygets ”performance” nämnvärt. Rapportfigurer och -uppställningar ska inte förändras till det sämre av att dessa blir fler.
- 5) **Klientserver-arkitektur** – OLAP-systemet ska stödja klientserver-miljö.
- 6) **Generisk dimensionalitet**
- 7) **Dynamisk ”sparse matrix”-hantering** – Mer eller mindre tomma tabeller – hur kan man presentera resultat utan dessa?
- 8) **Multianvändarstöd** – flera användare ska kunna hämta data ur systemet samtidigt.
- 9) **Obegränsad kalkyleringskapacitet** – beräkningar skall kunna göras kors och tvärs mellan dimensioner.
- 10) **Datamanipulation** - ”slicing-and-dicing”, ”drill-down” och andra manipulationer ska kunna göras genom att ”peka-och-klicka” och ”dra-och-släppa”.

11) Flexibel rapportering – förmågan att arrangera rader, celler och kolumner på ett sätt som underlättar analys genom visuell presentation av analytiska rapporter måste finnas.

12) Obegränsade dimensioner och aggregationer – man ska kunna ha många nivåer och dimensioner av data.

Sedan dessa regler presenterades har fler regler efterfrågats, samt att de regler som redan finns skall förfinas och utvecklas (Connolly & Begg, 1999).

2.3 Slut användare och datalager

För slut användare av datalager finns det problem av olika slag. Vi delar här in dem i följande kategorier: **Medverkansbrister** (vi syftar här på brister när det gäller användarnas medverkan i datalagerprojekt), **Bristfälligt användarstöd** (vi syftar här på behovet av hjälp från tekniker, eller det inbyggda stödet i verktyget), **Datautvinningsproblem** (vi syftar här på svårigheter i samband med sökning) och **Datapresentationsproblem** (hur ser det man får ut ur datalagret ut och går det att tolka?). Av dessa kategorier kan vi notera, att de två första mera kan ses som **orsaker** till sådana symptom som märks när man skall tillämpa datalagret (vi återkopplar till vår forskningsfråga (kap. 1.2), som både skall utreda **vad** för sorts problem och **varför!**).

2.3.1 Medverkansbrister

Enligt Söderström är det att rekommendera när man utvecklar ett datalager, att slut användaren finns med i seminarier som rör datalagret och ställer krav. Det är ju slut användaren som senare kommer att använda systemet - man kan säga att systemet måste anpassas till den som ska använda det! Detta eftersom ett datalager utnyttjas effektivast om beslutsfattarna själva kan använda det. Att använda systemet själv klarar man, om datastrukturen stämmer överens med ens (beslutsfattarens) egen syn på verksamheten. Det är även viktigt att i ett datalager lagra detaljer; genom att ha mycket detaljer blir det lättare att analysera och fatta bättre beslut.

Det finns flera olika **faktorer som bör styra valet av klientverktyg** (Söderström, 1997):

Anpassning till användaren:

- Hur generella gränssnitt klarar användaren av?
- Vilket är behovet av att kunna skapa anpassade dialogformulär?
- Vilken grad av hjälp måste användaren ha?
- Hur lång tid kan användaren spendera på utbildning?
- Är användaren en specialist, som kommer att ägna stor del av sin tid att arbeta med verktyget, eller kommer kontakten med verktyget att vara sporadisk?

Vilken är tillämpningen?

- Kommer användarna överhuvudtaget själva välja utdata?
- Hur snabbt behöver användaren få svar?
- Hur stor är variabiliteten i de typer av frågor som kommer att ställas?

Vilka är kraven på datatransport?

- Hur stora datavolymer behöver genomsökas för att utföra aktuella sammanställningar?
- Vilka typer av strukturer behöver verktyget kunna hantera?

Bischoff & Alexander (1997) menar, att användarnas medverkan är mycket viktig i alla faser av design och utveckling av datalagerlösningen. De ska delta i utformandet av kravspecifikationen, i datadefinieringen, i valet av slutanvändarverktyg, i utvecklingen av en kundanpassad applikation, i uppföljningen av designdokumentation, i designen av procedurer för att komma åt och manipulera data, i test och validering, i framtagningen av processer för hantering av förändrade krav, samt vid mätningen av användartillfredsställelse med lösningen.

En korrekt balans mellan inblandningen av IT-avdelningen och användarskaran är avgörande för slutresultatet (Bischoff & Alexander, 1997).

Bischoff & Alexander (1997) säger vidare, att det finns **sju ”dödssynder”** man ska undvika att falla på när man utvecklar en datalagerlösning (aktar man sig inte för dessa, blir slutresultatet inte så bra!):

- 1) **att** tro att ett informationsbehov kommer att uppstå när väl datalagret är skapat (att alltså inte använda redan formulerade krav som drivkraft),
- 2) **att missa** att skapa en arkitektur för datalagerlösningen att guida den fortsatta utvecklingen,
- 3) **att inte** inse vikten av att tidigt dokumentera alla vägval och ställningstaganden (ex ”hur mycket data skall laddas in i lagret?”),
- 4) **att** misshandera slutanvändarverktyg och tillgängliga metodologier för framtagningen av dessa,
- 5) **att inte** följa någon utvecklingscykel med faser under projektet,
- 6) **att inte** på ett korrekt sätt lösa ”datakonflikter” som kan ge inkonsistens i databasen (ex att olika avdelningar har olika format på data och det är svårt att enas om en standard), och
- 7) **att inte** kunna ta till sig av det som gått mindre bra och lära av misstag (om detta görs, kan misstagen guida en i hur felet ej skall upprepas).

Vi tycker att **synd nummer fyra** - misshandling av användarverktyg – är särskilt intressant för vårt arbete. Man måste nämligen ha en speciell **metodologi** vid framtagningen av slutanvändarverktyg. **Denna metodologi innefattar följande steg:** Bestämna urvalskriterier; ta reda vilka sorters verktyg som finns på marknaden som

uppfyller dessa kriterier; välja det bästa verktyget i varje kategori; hålla produktpresentationer och -demonstrationer; skapa prototyper, därefter välja det bästa verktyget och slutligen titta på de implementationer man redan har och bestämma till vilken grad dessa går att införliva med verktyget.

Anledningen till att sådana här punkter ställs upp tycks vara, att det är lätt att man annars inte gör detta arbete - Bischoff och Alexander och övriga skribenter i boken baserar sina råd på erfarenheter från en lång branschvana. Hortell & Soini (1999) har gjort iakttagelsen, att hänsyn till användares önskemål och medverkan i utvecklingsarbetet samt vid valet av verktyg saknas helt hos undersökta företag.

2.3.2 Bristfälligt användarstöd

När man utvecklar ett datalager, bör man redan från början ha stöd och dokumentation för att så snabbt som möjligt få igång datalagret.

Själva syftet med ett datalager är, att individerna själva ska kunna använda det. Det är därför, mer än vanliga system, beroende av individernas acceptans. Speciellt viktiga områden blir då utbildning och användarstöd. Slutanvändaren är den som ska trycka på knapparna och helst inte vara i behov av mellanhänder (Söderström, 1997). Båda rapporterna vi läst talar om hur användarna efter en kort utbildning får sitta och ”upptäcka själva”, samt lära varandra/av varandra.

Det är även viktigt att användarna har tillgång till så kallad metadata, som de själva kan förstå utan att behöva få hjälp av andra eller av någon specialist. En användare som inte har tillgång till metadata, kan inte utan hjälp göra analyser. Användare av datalager arbetar vanligen med fördefinierade frågor, rapporter och analyser - med eller utan parametrar. Dessa personer ägnar i allmänhet lite tid åt att förstå de möjligheter ett datalager faktiskt erbjuder. Användare behöver få tillgång till **bibliotek** med färdiga analyser, frågor och tillämpningar (Söderström, 1997).

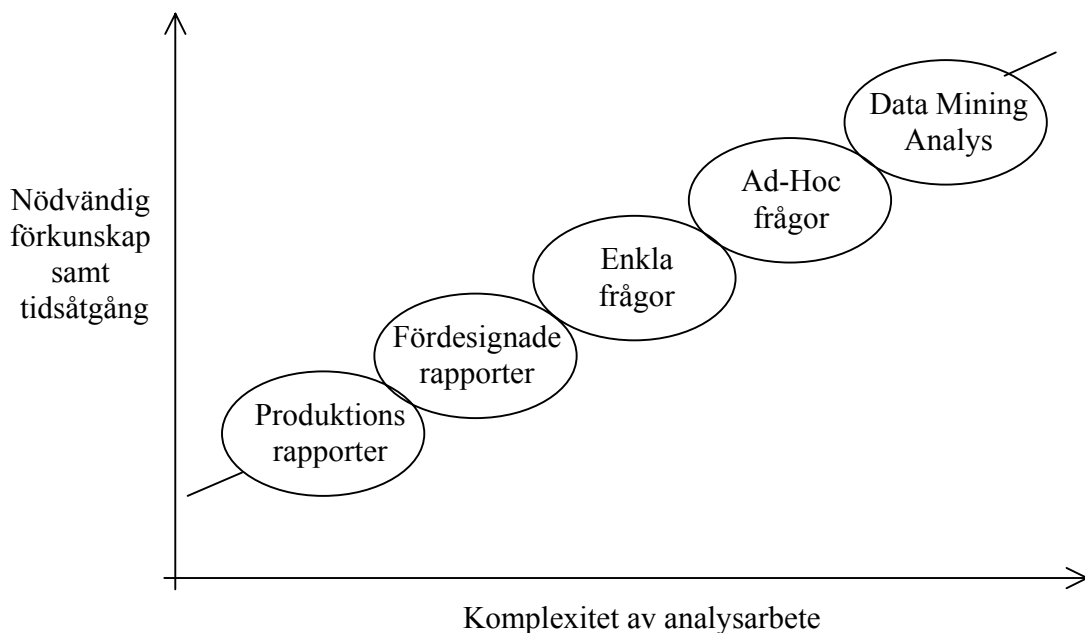
På grund av detta, menar Söderström (1997) vidare, är det viktigt att klientmiljön tillhandahåller följande:

- Generell hjälpfunktion för verktyget
- All dokumentation tillgänglig som en del av verktyget
- Åtkomst till metadata
- Data i dimensionstabeller, som inte anger koder utan beskrivande namn.

En annan viktig faktor är, att ha ett bra, enkelt och lättförståeligt **användargränssnitt**. Dessa gränssnitt ska vara **anpassade till användaren**. Data i ett datalager ska i första hand struktureras på ett för användaren begripligt sätt (Söderström, 1997). Det förekommer idag (Holmgård & Fazeli, 2003), att användare tycker den presenterade informationen är alltför detaljerad och att detta gör den svårtolkad.

Vilket klientverktyg man än väljer, bör valet så mycket som möjligt styras av egenskaper och behov hos användarna; att flytta, kontrollera, korrigera, transformera och utföra beräkningar på data från källsystemen tillhör de stora och svåra uppgifterna i ett datalager (Söderström, 1997). Om databasen till exempel nås via ett generellt gränssnitt (t ex ODBC), kan varje användargrupp i princip välja det verktyg som passar bäst för just dem.

Bilden nedan visar hur **kraven på förkunskaper och tid att ägna ökar**, beroende på hur komplext det analysarbete man skall utföra är:



*Figur 4: Ökande krav/komplexitet av analys.
Källa: omarbetad bild från Next Action Technology (1997).*

Den är hämtad från ett ”white paper”, kallat ”Filling the Fact Gap”, av det brittiska IT-företaget Next Action Technology (1997). Detta tar - förutom ökade användarkrav - upp det välkända så kallade **faktagapet**, d v s att företagsledare som har kunskap om branschen inte har kunskap om datasystemet, samtidigt som teknisk personal som kan utforma frågorna inte har kunskap om verksamheten. Detta fenomen återkommer i den mesta litteraturen och medför att datalagret inte kan tillämpas optimalt.

Next Action Technology (1997) säger vidare om vikten av utbildade användare, att utan analytisk kunskap kan så kallade ad-hoc-verktyg producera osäkra resultat. Frågor som dessa är inte ovanliga: Jag är inte säker på vad jag frågade efter! Hur många objekt kommer jag att få som svar? Har jag ställt en motsägelsefull fråga? Är frågan tillräckligt specificerad? eller; Hur kan jag kontrollera att resultatet stämmer?

Enligt Söderström (1997), är det även viktigt att fråga sig vilken kvalitet data har. Idag finns olika funktioner för att garantera att data är korrekt och tillgänglig. Det är viktigt att skapa ett förtroende för den information som hämtas ur datalagret - ofta är användarna skeptiska till den information de får ut och detta med rätta! Ansvaret för den eller de personer som ska ta hand om datakvaliteten skall definieras och användarna ska löpande informeras om kvalitén. Ett exempel på att denna riktlinje ofta förbises är, att flera **tillfrågade användare känner viss skepticism** för den data deras sökning genererat (Holmgård & Fazeli, 2003).

För att användare så självständigt som möjligt ska kunna använda datalagret och göra analyser, är det av vikt att krav och lösningar från utvecklingsarbetet dokumenterats. Ett sätt att göra detta är att skapa ett så kallat metadata-directory, som kan användas av både användare och tekniker (Alur & White, 1996). Metadata är data om data, d v s anger varje dataobjekts entydiga betydelse samt ursprung. I directoriet anges även om några förändringar av data gjorts och i sådana fall vilka. Användaren använder directoriet bland annat för att få reda på hur och var data är lagrat i datalagret - detta är nödvändigt för att kunna göra självständiga sökningar. Här får man också veta i vilka tabeller data förekommer, indexeringar, vyer, summeringar, aggregationer och beskrivning av hur aggregationer ska göras.

Man kan jämföra en datalageranvändare som saknar metadata-directory med en biblioteksbesökare som inte har någon kunskap om hur böcker ordnats och förvaras på biblioteket (Alur & White, 1996). En person som är bekant med bibliotekets sorterings-system och vet titeln på boken, kan gå direkt och söka upp den utan att behöva befatta sig med bibliotekets katalogiseringsystem och liknande. Detsamma gäller för den slutanvändare av datalager som är bekant med alla dataobjekts former och kategorier – han kan direkt lägga in rätt dimensioner i ”sökubnen”, utan att behöva befatta sig med tekniska aspekter av datalagret.

Metadata skall även lagras så att frågehantaren kan använda sig av den för att exempelvis styra rätt fråga till rätt tabeller. I stort sett alla delar av datalagermiljön är på något sätt beroende av metadata för sin funktion (Connolly & Begg, 1999).

2.3.3 Datautvinningsproblem

Det svenska företaget Business Navigator talar i sin tekniska rapport ”Datalagerenklare och billigare” (1998) om tre huvudfaser i datalagerutvecklingen. De två första faserna handlar om hur man tar fram själva databasen, medan den tredje syftar till att göra innehållet i datalagret tillgängligt för slutanvändarna. Man har funnit att det är vanligt att företag lägger mycket energi, kunskap och tid på de två första faserna, medan den sista inte ses som lika viktig. Teknisk kunskap tenderar att totalt dominera diskussionen om datalager, medan användarproblem inte diskuteras på samma sätt. Följden av detta blir, att användare har svårt att använda de verktyg som tillhandahålls - de saknar både vana och utbildning att tillämpa analyser och data på rätt sätt och systemet ger inte svar på deras frågeställningar. Business Navigator (1998) säger också, att värdet av investeringen i ett datalager står och faller med frågeverktyget.

Det är likvärd av betydelse för slutanvändaren vilken design man använder på databasen. Det finns två huvudstrukturer att välja mellan – stjärn- eller snöflingestruktur (se begreppsdefinition). Det är i första hand söktiderna man vill korta ner genom att välja rätt struktur (med olika strukturer måste olika datamängder gås igenom vid sökning).

2.3.4 Datapresentationsproblem

Enligt Söderström (1997), finns det olika användargränssnitt man kan använda sig av för att presentera information:

- Färdiga rapporter, som då saknar parametrar
- Färdiga formulär, som bara har ett fåtal parametrar och därför är lätta att använda
- Färdiga frågor, d v s återutnyttjande av en analys som någon annan gjort i ett klientverktyg direkt mot en generell struktur
- Att utnyttja ett klientverktyg själv direkt mot en generell struktur

I de två sistnämnda fallen arbetar användarna direkt mot en datastruktur, som måste utformas efter deras behov. Oftast används **stjärnstruktur** eller **snöflingescema**.

Datalager innebär ofta stora datamängder. Även om parallell teknik och snabba datorer underlättar processandet, behövs ytterligare designval såsom aggregering, partitionering, kodade attribut och indexering för att förbättra prestandan. Det gäller att ta ställning till hur länge data ska lagras i datalagret, vanligen är fem år (Söderström, 1997). Den stora utmaningen inför framtiden är, att kunna låta datalagret växa (då organisationer använder sig av allt större datamängder) utan att detta inverkar alltför mycket på svarstiderna. Redan idag anser många slutanvändare att svarstiderna för sökningar är för långa. Det är samtidigt en fråga om vad man kan förvänta sig; användare utan teknisk kunskap måste bli upplysta om vad det är rimligt att vänta sig för svarstid på exempelvis en sökning som kräver genomgång av hela datalagret. Detta tar även Connolly & Begg (1999) upp – de menar, att ett dilemma i informationsprocessen är hur man ska kunna processa större och större datamängder utan att få för långa svarstider.

Sammanfattningsvis kan man säga, att teorins syn på vårt problemområde är att det finns en hel del användarproblem, som är tämligen vanligt förekommande och som man helst bör undvika genom att utveckla datalagerlösningen på rätt sätt - det är svårare att rätta till problemen när de väl uppstått. Likvärd återkommer samma typer av problem i de böcker vi läst (som är av så olika årgång!) - det har alltså inte hänt så mycket på det området under det senaste åren, vilket tyder på att problemen är svårlösta.

3 Metod

Här beskriver vi hur vi arbetat, steg för steg, för att uppnå vårt syfte och finna svar på forskningsfrågan.

3.1 Litteraturstudie

”Vi kan inte förstå något över huvudtaget utan förförståelse” (Thurén, 2002).

Den förförståelse vi hade utökade vi genom att fördjupa oss i teorin om ämnet. Detta gjorde vi genom att låna böcker på biblioteket, läsa rapporter som handlade om datalager, men eftersom alla böcker vi hittade var några år gamla valde vi att även söka på Internet efter artiklar som var nypublicerade. Genom att, på detta sätt, få en uppfattning om läget hos företag idag vad gäller datalageranvändarnas situation, fick vi ett bättre utgångsläge inför valet av intervjufrågor och intervjuobjekt, samt att utföra och förstå intervjuerna och analysera intervjuvaren

3.2 Undersökningsmetod

Då vår utredning var av förklarande/beskrivande natur, vårt syfte samt forskningsfråga ställdes så att undersökningen skulle komma att göras i här-och-nu-miljö och på pågående fenomen (ej historiska) och vi som utredare hade liten möjlighet att styra dessa fenomen på arbetsplatsen (Yin, 1994), föll det sig naturligt för oss att välja **fallstudie** med **kvalitativa intervjuer** som undersökningsmetod. Den kvalitativa metoden går ut på att visa komplexa sociala fenomen och betydelsefulla drag hos företeelser här och nu, så som organisationsprocesser och -mognad, där andra undersökningsmetoder inte räcker till. Den svarar på frågor som **hur** eller **varför** (Yin, 1994).

Intervjuer ger följande fördelar framför andra metoder: fullständigare svar, möjlighet att förklara och följa upp frågor, få tilläggsinformation, når rätt person samt ger utrymme för respondenter att berätta och beskriva. Fullständigare svar skulle vara en klar fördel för oss eftersom vår uppgift var att kartlägga olika svårigheter och att få tilläggsinformation är också värdefullt vid en kartläggning.

Det finns två olika sätt att genomföra intervjuer: standardisering (frågor bestämda i förväg) och strukturering (ordningen bestämd i förväg). Vi har valt att använda oss av en mellanvariant (som är den som oftast används) kallad **halvstrukturerad intervju** – man har frågorna i förväg, kommer med följdfrågor, tar bort det man redan fått svar på och kan byta ordning på frågorna.

Det passade oss även att införliva vissa **kvantitativa** moment med intervjuerna - detta gjorde vi i form av enkäter, med fasta frågor/svarsalternativ. Den kvantitativa

undersökningsmetoden svarar på frågor som **vem, var, hur många** eller **hur mycket** (Yin, 1994).

Enkäter ger följande fördelar framför andra undersökningsmetoder: billigare och, lätt att nå fler personer, geografisk spridning, ingen intervjuareffekt, mindre prestige och känslighetshinder samt ger möjlighet att tänka igenom svaren. Då vi velat få geografisk spridning på vår undersökning var enkäter särskilt lämpligt och då våra intervjufrågor kunde ha haft känslighetshinder var enkäter bra att börja med.

3.2.1 Datainsamling

Huvudsyftet med enkäterna var att få fram lämpliga intervjuobjekt. Vi valde att skicka ut dessa till företag som är stora, rikstäckande, vinstdrivande, fattar strategiska beslut med hjälp av datalagret, är välkända, samt ägnar sig åt helt olika affärsverksamhet. (Genom detta ville visa, att samma typer av användarproblem kan förekomma hos helt olika företag). Vi skickade ut enkäten i form av ett brev eller via e-mail till tjugo företag, varav åtta sedan svarade.

När vi fått in tillräckligt många svar och sett vilka företag som hade användarproblem, valde vi ut några företag lämpliga för intervju. Kravet på våra respondenter var, att de skulle ha sagt sig ha någon typ av problem i datalageranvändningen, eftersom endast dessa var intressanta för vår del, samt angett intresse för vidare kontakt. Vi valde att utföra intervjuerna med en beslutsfattare eller en tekniker (som tekniker klassar vi de personer som utvecklat datalagerlösningen åt företaget) på respektive företag, och utelämnade annan personal. Vi överlät åt företagen att själva välja vem på företaget som skulle delta i vår undersökning - detta eftersom våra frågor var av sådan karaktär, att såväl användare som tekniker uppriktigt skulle kunna svara på dem. Det blev samma personer som svarat på enkäterna.

Vi utförde intervjuerna på fyra olika företag och pratade med totalt sex personer. Eftersom bara ett av det företag vi valt ut till undersökningen fanns här i Luleå, blev det tre telefonintervjuer (en på plats). Vi spelade in alla intervjuerna, för att inte missa något viktigt.

Vi kom också att använda enkätsvaren i analysen av materialet.

3.3 Analysmetod

Vi valde att analysera materialet i två steg; först svaren från den enkät vi skickade ut för att kunna välja intervjuobjekt och sedan resultaten från intervjuerna. Enkätsvaren jämfördes sinsemellan och intervju svaren både med varandra och med relevant litteratur.

Det finns två huvudstrategier att välja mellan när man bestämmer analysmetod; antingen förlitar man sig på det som står i litteraturen och väger in den i analysen, eller så gör man en helt fristående fallstudie utan inverkan av litteraturen (Yin, 1994). Vår

undersökning föll under den första av dessa. Dessutom har vi analyserat vårt material enligt en modell som kallas mönstermatchning. Den går ut på att man jämför de mönster man kan utläsa av undersökningsresultatet med redan uppställda mönster (Yin, 1994).

När vi hade fått in allt material och sammanställt svaren i tabeller, var det tid att börja jämföra de svar vi fått genom intervjuerna. Vi jämförde de olika företagssvaren med varandra, för att se om några företag svarat lika på några frågor. Sedan jämförde vi svaren mot de föreskrifter vi haft i teorin, om bland annat utbildning för användare, vad verktygen skulle klara av, samt användarnas medverkan i hela processen för att ta fram datalagerlösningen.

Som ett sista steg i forskningsarbetet var det dags att börja dra några slutsatser - fanns några kopplingar, stämde våra teorier och det vi själva trott? Vi tänkte, att vårt bidrag till forskningen skulle bli att visa på att det finns brister hos verktygen, eller andra svårigheter i användandet av dessa.

3.4 Validitet och Reliabilitet

Validitet hos en rapport har att göra med hurvida man verkligen undersökt det man föresatt sig att undersöka. Man måste till exempel ha ställt rätt frågor vid undersökningen, så att man får fram rätt data för att kunna dra några slutsatser.

Vi valde att göra både en kvantitativ och en kvalitativ undersökning, eftersom ingen av dem skulle varit tillräckliga var för sig. Genom den kvantitativa enkäten fick vi tidigt se vilka problem som förekommer hos företag – något som haft betydelse för valet av respondenter och slutsatsdragningen. De kvalitativa intervjuerna har försett oss med data som man inte får fram enbart genom enkäter, samt att det under dessa även framkommit ytterligare problem som ej nämndes i enkätsvaren. Att bara göra en litteraturstudie hade visserligen visat på en del problem - men långt ifrån alla - och de problem som man hade fått fram skulle inte ha varit representativa för Sverige, eftersom den mesta litteraturen är utländsk. Genom att bara utföra en fallstudie utan litteratur hade vi fått mycket sämre underlag för frågeställningar, eftersom vi grundat våra frågor på litteraturen. Validiteten hade då blivit sämre.

Vi tycker att validiteten på vår rapport är god - intervjufrågorna har baserats på det litteraturen sagt om förekommande problem, vi har vänt oss till dem som verkligen arbetar påtagligt med de problem vi ville studera och vi har fått sådana svar vid intervjuerna att vi kunnat dra slutsatser och se likheter. (Vi har därmed uppnått syftet med vårt arbete.)

Reliabilitet innebär tillförlitlighet till resultatet – stämmer det och är det representativt?

Att utföra intervjuer på plats är det bästa sättet, men då våra respondenter fanns i andra delar av landet blev det telefonintervjuer. Vi bandade dem dock - därför förlorade vi

nog inte så mycket på detta (det är annars vanligt att man glömmer sagda saker vid intervjun). Vi tror därför inte att vår reliabilitet påverkats negativt av telefonintervjuer.

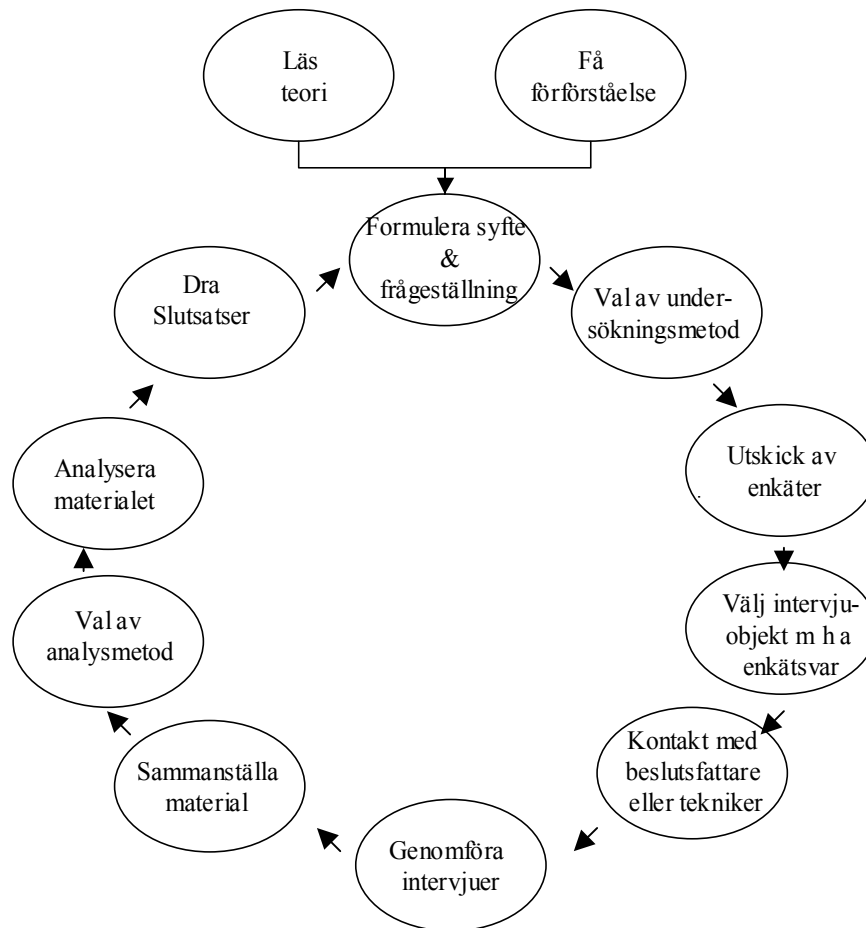
De artiklar/utdrag ur artiklar från Internet som vi använt oss av är skrivna av professionella systemutvecklare eller tillverkare av verktyg för att lösa användarproblem, i **syfte att dela med sig av kunskaper och erfarenheter** - vi tycker därför att de har hög tillförlitlighet. (Dessa har också varit mycket nyare än de böcker vi använt.)

Vi skickade ut våra intervjufrågor i förväg, för att ge respondenterna tid att fundera på dem. Om vi hade använt oss spontana frågor vid intervjuerna, skulle både respondenterna och vi ha missat viktiga frågor och svar. Reliabiliteten hade på detta sätt blivit klart sämre.

Man kan alltid ifrågasätta uppriktigheten hos de som deltar i kvalitativa intervjuer; har någon av dem något motiv för att hellre vilja ge oss de svar man tror att vi "vill ha" än att vara helt ärlig, eller finns någon med som har anledning att vilja vinkla svaren åt så positivt håll som möjligt för honom/henne..? I detta fall kunde man ifrågasätta om teknikerna i undersökningen ger oss lika uppriktiga och uttömmande svar som användarna – har de varit rädda för att framstå som sämre IT-personal om de uppger alltför många indikerade problem från "sina" användare..? Vi tror inte det. Vi frågar dem om bekanta, erkänt svårlösta problem som ingen har någon ultimata lösning på, i syfte att skapa medvetande och sträva efter allmängiltiga lösningar. Det tillmötesgående och intresse vi fått av såväl användare som tekniker i samband med intervjuerna, förstärker vår inställning.

Vi tror att reliabiliteten på vår rapport är god, inte minst eftersom de problem som framkommit i fallstudien stämmer med vad som sagts i den litteratur vi läst. Sannolikheten att en annan grupp, som väljer ut fyra nya företag, skulle uppnå samma resultat är enligt oss stor.

3.5 Sammanfattning av vårt arbete:



Figur 5: Metodmodell.

Källa: Egen

(Bilden skall uppfattas som en iterativ rotation genom olika ”arbetsstationer”; när vi dragit de första slutsatserna, påbörjas ett nytt varv, där vi bygger på, stämmer av och omformulerar.)

4 Empiri

4.1 Enkätssammanställning

Vi väljer att ta hänsyn till enkätsvaren inte bara från de företag vi intervjuat, utan även från andra som svarat, men som inte ingått i fallstudien. Detta för att kartlägga om det finns några framträdande problem genom att flera svarat ungefär det samma, samt få stöd för de slutsatser vi drar av intervju svaren. Samtliga företag vi presenterar i enkätssammanställningen säljer varor eller tjänster och använder sitt datalager till strategiskt beslutsfattande. De är alla stora och väletablerade på den svenska marknaden. Då vi inte uttalat bett om företagets tillåtelse att avslöja deras namn, kallar vi dem här A, B, o s v. Företagets namn har här inte heller någon större betydelse för att kunna kartlägga förekommande användarproblem. Vi redovisar här alla enkätfrågor utom den första (Använder Ni datalager?), som alla svarat ja på - en förutsättning för att vi skulle ta med företagen i undersökningen. Företagen har själva fått utse en person (slutanvändare eller tekniker) de ansett lämplig att svara på frågorna. Nedan visar vi enkätsvaren från samtliga företag som svarat; här ingår alltså även de företag som vi senare valde att intervju (F, G, E, resp. A – jfr kommande företagsbeskrivning).

Sammanställning av enkät svar:

	A	B	C	D	E	F	G
<i>Hur länge har ni använt datalager?</i>	5 år	2 år inom Point-of-sales data	2 år	6 år	3 år	2 år	3 år
<i>Hanterar & underhåller ni ert datalager själv?</i>	Ja	Ja	Både egen och inhyrd personal	Ja	Ja	Hyr in	Ja
<i>Problem?</i>	Långsamt, stora och komplicerade frågor ställs. Svajande performance.	Problem med stora data-mängder, inläsnings-apparat, behov av olika data, samt definitions-problem.	I början var svars-tiderna för långa, men de är nu OK. Användar-gränssnittet uppfattades heller inte särskilt bra av vissa användare. Därför har man skapat ett eget enklare gränssnitt.	Data kvalitet. Tolkning av information. Har upplevt problem med svars-tider och svåra sökningar.	Att få tredjeparts-rapport-och-analys-verktyget att fungera bra mot datalagret.	Långa svarstider p g a stora mängder data. Begränsningar av sökningar vid val av många variabler.	Det har varit problem att få tag i data. Förståelse av DW-konceptet från vissa länder. Vissa problem med integrations-miljön som har orsakat avstämnings-problem.

Tabell 1: Enkätssammanställning

4.1.1 Företags- och respondentbeskrivning – intervjuobjekt

Stadium

Stadiumidén tog fart 1974, när Ulf Eklöf tog över den lilla sportbutiken Spiralen Sport i Norrköping. Visionen då var att skapa ett sportvaruhus och göra sportprodukter tillgängliga för alla. 1983 var Spiralen Sport Sveriges största sportbutik.

Stadiumbutiken öppnades 1987 i Stockholm. Stadium blev en arena för sport och fritid. Stadiumidén växte och fler butiker öppnades över hela landet. I dag finns totalt ca 80 Stadiumbutiker i Sverige, Danmark och Finland och visionen om att inspirera till en aktiv, rolig och hälsosam vardag för många människor lever vidare. Stadiumbutikerna ägs och drivs av moderbolaget Stadium AB, som också är servicekontor för butikerna. I Stadium AB ingår också ett tiotal andra verksamheter och butiker.

Stadium använder sig av datalager sedan två år tillbaka, för uppföljning, prognoser, inköp, försäljning och kunddata. Man hyr in personal från konsultföretag för förvaltning och support av datalagret. Datalagret fungerar överlag bra, men vissa problem förekommer.

www.stadium.se

Respondentbeskrivning: Respondent A arbetar på Stadiums marknadsavdelning, och använder sig av datalager för CRM (Analyser på kund- och produktnivå). A är alltså en **slutanvändare** av datalager. Analysverktyget är ProClarity – ett OLAP-verktyg.

Findus

Findus Sverige AB är ett av landets ledande livsmedelsföretag. Djupfryst utgör den största delen av deras försäljning, med allt från färdigrätter och fisk till ärter och wok i sortimentet. På den svenska marknaden säljer de även kolonialprodukter, bland annat majonnäs, soppor, pastasås och marmelad. Findus är ett globalt företag med representation över hela världen.

Findus produkter finns i dagligvarubutiker över hela landet, men de gör även mat som serveras på till exempel restauranger, skolor och sjukhus. Findus Sverige AB har drygt 1400 medarbetare. Det svenska kontoret ligger i Bjuv utanför Helsingborg. Här sker en stor del av produktionen, men företaget har också fabriker i Helsingborg och Loftahammar. Hög och jämn kvalitet är den viktigaste ingrediensen i alla Findus produkter och höga kvalitetskrav är en trygghet för konsumenterna.

Findus har använt datalager i tre år för produktions- och säljutveckling. Historiska data används för framtidsbeslut om nyutveckling och produktlansering (vinna marknadsandelar). Datalagret fungerar överlag bra, men vissa problem förekommer.

www.findus.se

Respondentbeskrivning: Respondent B och C har varit med och tagit fram Findus datalager, de är med andra ord **tekniker** i vår undersökning. Man sköter och underhåller datalagret själva, genom Findus egen dataavdelning.

Vattenfall

Redan 1909 började Vattenfall att producera och leverera elektricitet i Sverige. Idag finns deras marknad i stora delar av Europa. Vattenfall lägger stor vikt vid att utveckla tjänster och erbjudanden för sina kunder. El är nog det första man tänker på när man hör Vattenfall, men Vattenfall är mer än så, mobiltelefoni, till exempel. Vattenfall erbjuder en rad olika tjänster och produkter på energimarknaderna i Sverige och Europa. Vattenfalls kunder kan själva välja energislag från ett brett utbud av miljöanpassade produkter.

Vattenfall Datas roll i Vattenfall är att vara ett Shared Service Center (SSC) vilket innebär att Vattenfall Data är ett stödbolag som styrs mot kostnadstäckning. Resultatet av Vattenfall Datas insats återfinns i Business Units affär. Vattenfall Data är ett eget aktiebolag med mål och avkastningskrav som fastställs av bolagets styrelse. Deras vision är att i närhet och samspel med koncernens enheter skapa ett världsledande internt IT-bolag. Deras affärsidé är att med Norden som hemmamarknad förstärka Vattenfalls konkurrenskraft med efterfrågade och effektiva IS/IT-lösningar. Vattenfall Data har 350 anställda.

Slutanvändarna på Vattenfall Data fattar inga direkta beslut m h a datalagret utan enbart till att presentera och rapportera information för andra på Vattenfall. Man har använt sig av datalager i tre år. Man underhåller och sköter datalagret själva med visst stöd från en fristående extern konsult. Vattenfall Data har en datlagerlösning som heter SAP BW, där samma leverantör stått för både datakällan och slutanvändarverktyget. Datalagret fungerar överlag bra men vissa problem förekommer.

www.vattenfall.se

Respondentbeskrivning: Respondent D och E arbetar båda på Vattenfall Data. Respondent D har hand om underhållet av datalagret och ger support åt användare av datalagret. (Han har arbetat med detta i ett halvår ungefär.) Vi klassar honom som **tekniker**. E är sedan ett år tillbaka användare av datalagret på så sätt att han tar fram rapporter som beslutsunderlag åt beslutsfattare. Han anser sig därför inte vara slutanvändare i ordets rätta bemärkelse (han gör inga dataanalyser). Det är svårt att sätta någon särskild kategori på honom, då han samtidigt ger support åt användare i form av att ta fram beslutsunderlag åt dem - vi får därför klassa honom som ett **mellanting mellan slutanvändare och tekniker**.

Vattenfall Data använder sig av koncernkonsolideringssystemet Frango, via ett egenbyggt överföringsprogram (laddhanterare) till datalagret (SAP BW) och till rapportverktyget BusinessObjects (frågehanterare).

Frontec

Frontec är ett börsnoterat IT-konsultföretag, med fokus på konsulttjänster inom eBusiness. Man är specialister på integration. Frontec hjälper kunder att få processer och IT-stöd att samverka.

Sedan starten 1981 har Frontec utvecklats till ett av de ledande IT-konsultföretagen i Sverige. De finns representerade på 8 orter i Sverige, från Luleå i norr till Malmö i söder, med ca 500 medarbetare. Frontecs verksamhet inom eBusiness finns idag samlad i fem regioner, med fokus på den lokala marknaden. Dessutom finns ett nationellt utbud representerat av enheterna SAP, Meridium och Multidesign. Meridium fokuserar på systemutveckling med Microsoft-produkter och SAP arbetar med affärssystem genom implementation av SAP R3. Frontecs vision är att utveckla kundernas verksamhet mot eBusiness, genom att vara specialister på integration av processer och teknik och med ett arbetssätt som enklare och snabbare ger resultat.

Affärsidén gäller för hela verksamheten och syftar till att åstadkomma kundnytta genom att erbjuda ett stort utbud av tjänster och IT-lösningar, där alla delar i kundernas verksamhet - organisation, processer och teknologi - integreras. En lösning från Frontec ska alltid korta ledtider i kundens verksamhet bättre än lösningar från andra leverantörer. Genom att effektivare genomföra innovativa åtgärder med hög kvalitet, ska man spara tid åt kunderna och förbättra deras möjligheter till affärer nu och i framtiden.

www.frontec.se

Respondentbeskrivning: Respondent F arbetar som konsult åt Frontec och svarar därmed utifrån sina erfarenheter ute hos kundföretagen. F lämnar inga svar som slutanvändare för ett visst företag, utan svarar allmänt som **tekniker**. Hon har arbetat som konsult sedan 1995. F anser att datalagret oftast fungerar bra, men vissa problem förekommer.

4.2 Sammanställning av intervjuvaren

Här sammanställer vi intervjuvaren i sin helhet, redovisade fråga för fråga. Vi har för överskådliggighets skull kategoriserat frågorna enligt samma teman vi använder i Teori.

Medverkansbrister

	Stadium	Findus	Vattenfall	Frontec
<i>Har du fått vara med och bestämma inköp av verktyg?</i>	Användarna fick välja mellan två redan bestämda verktyg. Valet av verktyg styrdes av de verktyg som redan fanns i företaget.	En användargrupp fanns med vid bestämmandet av vilket verktyg som skulle inköpas. Man har erfarenhet av att användare i någon form bör finnas med så tidigt som möjligt. Användarnas krav styrde inköp av verktyg.	Inga användare fanns med vid bestämmandet av inköp av verktyg, detta för att verktygsleverantörens produkter redan fanns på företaget och man valde då deras verktyg till datalagret.	Oftast har företaget redan en uppfattning om vilket verktyg man ska välja (ex. är "inne"), därför tillfrågas inte slutanvändare. Valet av verktyg styrs inte av behovet utan behoven styrs snarare av verktygen.
<i>Har du deltagit i utveckling av datalagret?</i>	Ja, redan från början.	Har en internationell arbetsgrupp/styrgrupp som består av både IT-personal och användare. Användarna har således representanter med i beslutandeprocessen.	Dataavdelningen har gjort en lösning som tillgodoser Vattenfalls rapporteringsbehov. För datalagrets ändamål har man inte tyckt sig behöva så mycket synpunkter från slutanvändare.	De har de oftast inte eftersom datalager projekten oftast är teknikdrivna istället för informationsdrivna, d v s enskilda personer på företaget är teknikintresserade och därför är inte användarnas krav i centrum.
<i>Är verktyget anpassat efter dig?</i>	Ej efter varje specifik användare, men däremot används fördesignade rapporter.	Ej efter varje specifik användare, men däremot används fördesignade rapporter.	Ej efter varje specifik användare, men rapportbehovet tillgodoses på ett bra sätt genom så kallade datamarts (man behöver då inte gå igenom hela datalagret utan bara valda delar av det).	Den vanligaste kundanpassningen är genom att man tar redan på vilka sökvägar användaren oftast använder och anpassar gränssnitten efter dessa.

Tabell 2: Medverkansbrister

Bristfälligt användarstöd

	Stadium	Findus	Vattenfall	Frontec
<i>Klarar du av att göra sökningarna själv?</i>	Användarna behärskar inte att konstruera kuberna själva eftersom de saknar riktig kunskap om hur ett datalager fungerar. Därför används för designade kuber utformade av IT-avdelningen.	Man använder sig av fördesignade kuber och dessa klarar användarna att hantera. Standardiserade webbaserade rapporter används. Användarna har befriats från att göra ad-hoc-sökningar genom att man istället lagt in en ad-in-funktion mot OLAP-databasen i Excel.	Eftersom inga fullständiga sökningar görs av slutanvändare blir denna fråga lite svår att svara på.	Varierar beroende på användare, sofistikerade användare kan göra kuberna själva, medan mindre kunniga användare oftast har fördesignade kuber, det går även att beställa sökningar. Man kan även få hjälp med att bedöma vilken information man behöver.
<i>Finns det inbyggt användarstöd?</i>	Någon form av hjälpdialog finns, däremot används den inte då det går lika snabbt att ringa och få svar av verktygsleverantören.	Någon form av användarstöd finns.	Användarstöd finns i form av support som man kan ringa till vid problem. Man har ett internt kompetenscenter för enklare frågor, detta kräver dock att verktyget används flitigt.	Verktygen har oftast användarstöd i form av metadata-directory detta räcker oftast långt som hjälp. Det finns oftast även ett telefonnummer man kan ringa eller/och e-mailadress.
<i>Har du fått någon utbildning för att använda verktygen?</i>	Kort intern utbildning därefter löpande kommunikation med konsulter från IT-avdelningen.	Utbildning i "steg" beroende på användarens förkunskaper. Användarna utbildar också varandra.	Kortare interna kurser i de applikationer data-lagerverktygen opererar i. Man träffar också regelbundet andra som har samma verktyg och ventilerar problem som har uppstått.	Vanligt med en kort utbildning på en till två timmar.
<i>Har man dokumenterat krav & lösningar så att nya användare kan börja använda systemet?</i>	Bristfällig dokumentation om lösningar och kravspecifikation finns. För en ny användare är det dock viktigare att lära sig principerna av analysverktyget. Detta kombinerat med kunskaper om vår kunddata är tillräckligt för en ny användare.	Metadata och begrepp har dokumenterats. Anser att det är mycket viktigt att dokumentera alla begrepp som används för nya användare, det behöver då aldrig bli någon diskussion om vad ett ord betyder.	Karv och lösningar är dokumenterade, verktyget har ett inbyggt metadata-directory där man lagt in sådant som ordens betydelse och liknande uppgifter.	Har erfarenhet av att samtliga krav som fanns dokumenterades. Man dokumenterade om kravet prioriteras eller inte, om det realiserats eller inte, och om det realiserats så dokumenterar man hur det görs. Detta kräver att man har någon form av kravhanteringsverktyg. Det har även förekommit att man samlat ihop krav och dokumenterat dessa, samt vilka krav man realiserat. Oftast har man gjort detta för att hålla reda på systemets krav och för att kunna använda det som primärmaterial. Inte så mycket för att användarna ska kunna komma in och börja använda det, utan snarast för att kunna säga att det här kan ni göra.

Tabell 3: Bristfälligt användarstöd

Datautvinningsproblem

	Stadium	Findus	Vattenfall	Frontec
<i>Har ni några problem med att använda datalagret? Vilka problem?</i>	Problem med presentation p g a stora mängder data. Långa svarstider då man hanterar stora mängder data, samt vissa begränsningar av sökningar vid val av många variabler. Datorerna gör emellanåt timeout då de blir överbelastade (har provat med att köpa mer minne men inte hjälpt).	Det har varit problem att få tag i data från vissa länder, man har helt enkelt sämre system där. Förståelse av datalager-konceptet från vissa länder. Vissa problem med integrationsmiljön som har orsakat avstämningsproblem. Ej riktigt medvetna om verktygets totala kapacitet. Man har inte fullständig kontroll på all information som skall/har strömmat in i datalagret. Kuber har visat sig vara för stora. Användarna har haft problem med att förstå det man får ut, detta har man nu löst.	Prestanda är det återkommande problemet vad gäller användning av datalagret. Hur data ska modelleras och lagras för att prestandan ska bli bra är en ständig fråga. Bristande intresse från användarnas sida. Eftersom man har olika "laddtider" skulle man behöva en kvalitetskontroll på att se att den data som finns är rätt och riktig. Hur många datakällor som används, detaljeringsnivå på data, struktur av data är andra aspekter. Att använda rapportverktyget i ett DW är dock inte så komplicerat, frågan är dock om datat stämmer. Mycket arbete med att få (tredje-parts-) rapport- och analysverktyget att fungera bra mot datalagret.	Man arbetar mycket med att få kortare svarstider och enklare sökningar och lyckas ganska väl. Men i vissa lägen kan det vara svårt att få data från levererande system. Det har även varit problem med bristande data-kvalitet. Även problem med tolkning av informationen förekommer. Tveksamhet om all information man behöver finns i datalagret och om informationen rätt och riktig. Långa svarstider är ofta ett problem och hur man ska veta om sökningen blir klar i tid. Problem med laddtider p g a stora mängder data. Systemens kapacitet är ofta för låg för att klara av att ladda in större mängder data tillräckligt ofta. Konflikter angående dataäggande (all data ska inte alla ha tillgång till)
<i>Vad tror du att dessa problem beror på?</i>	För stora datamängder gör att datorerna inte klarar av hanteringen smidigt. Problemen beror på mycket stora datamängder, kombinerat med för lite interminne i de datorer som används för analyser. Ställer man en fråga på lägsta nivå, orkar inte datorn processa detta och inget resultat kan presenteras.	Svårigheten att kunna leverera den informationen man behöver från vissa länder. Viss info saknas i deras lokala system, kvalitén kan ibland vara dålig och åtkomligheten kan vara svår.	Bristande intresse från användare kan bero på att datalagret modellerats felaktigt – det drabbar användaren i slutändan, genom att prestandan försämras. Tillförlitligheten till att den data man får ut stämmer kan förbättras genom kvalitetskontroller.	Politik inom organisationerna. Man är van att jobba i ett annat system. Det är ibland svårt att börja med något nytt. Datalager började som en teknikdriven "fluga", inte av krav från verksamheten.
<i>Hur tror du dessa problem kan lösas?</i>	Integrering av kuber, virtuella kuber (ej behöva processa två hela kuber utan bara vissa dimensioner av dem i en tredje kub). Man löser detta tillfälligt genom att köra analyserna direkt på servern och långsiktigt genom att skapa index för vissa frågor som ställs återkommande.	Problemen beroende på att flera länder är inblandade, kan kanske lösas genom att fortsätta utrullningen av de europeiska transaktions-system är den bästa lösningen på lång sikt. På kort sikt löses detta med stöd-system från centralt håll. För stora kuber har brutits ner i mindre delkuber, vilket gav kortare uppdateringstider, samt att backuper blev lättare att göra och tog mindre utrymme. Man har trimmat sin OLAP-server för att få kortare svarstider. Att användare inte förstår det de får ut datalagret har lösts genom fördesign & web-verktyg	Prestandan kan optimeras genom ändrade lagringsformer.	Långa svarstider kan lösas genom att ta reda på användarens vanligaste sökvägar och snabba upp dem. Problemet med dataäggande kan lösas genom att använda sig av så kallad datasteward, som handhåller data och avgör vem som har tillgång till vad. Visa vad man kan åstadkomma. Snabba prototyper som ger snabb kundnytta (användarnytta). Enkla användargränssnitt med stor flexibilitet.

Tabell 4: Datautvinningsproblem

Datapresentationsproblem

	Stadium	Findus	Vattenfall	Frontec
<i>Får du svaren i rimlig tid?</i>	Överlag bra svarstider, men vid större sökningar blir det väl så lång svarstid.	Har idag bra svarstider.	Långa svarstider förekommer då man har stora datamängder och beroende på hur datalagret är uppbyggt.	Långa svarstider rätt så vanligt, men ibland beroende på att slutanvändaren ställt sökfrågan fel.
<i>Är informationen presenterat i det format du vill ha det i?</i>	Ej direkt från verktyget, man använder sig av Excel för att få en mer lättolkad utskrift.	Ej direkt från verktyget, man använder sig oftast av Excel då de graferna som används är svår tolkade.	Eftersom användarnas uppgift är att vidarebefordra information till beslutsfattare, har man statiska rapporter som inte kan manipuleras. De användare som ändå vill göra det klipper in det i Excel.	Ofta är informationen inte i det format användaren önskar. T e x presenterar verktygen oftast sökresultatet i form av grafer vilket många tycker är svårtolkat, svårt att göra jämförelser samt svårt att spara, därför väljer många att använda sig av Excel.
<i>Är kvaliteten på det som presenteras bra?</i>	Alla beräkningar och rapporter är ännu inte helt kvalitetssäkrade. Men generellt sett är det bra kvalitét på dessa. Vi är som jag nämnde inne i en stor förändringsprocess på Stadium för närvarande. Ska byta kassasystem som kommer innebära att flödet av data och kvalitén på den samma ses över och förbättras.	Våra webbaserade standardrapporter anser vi vara lättolkade genom att de enkelt kan selektera via menyer. Samtliga rapporter har standardutseende.	Eftersom det är koncernredovisnings-siffror som visas, är det nödvändigt att det är korrekt. Kvaliteten är bra, förutom att några enstaka gånger har siffrorna blivit dubblade i slutanvändarnas rapporter. Orsaken till dubblade värden har varit det komplicerade dataflödet (transaktionsdata och styrtabelldata) från källsystemet	Det förekommer rätt så ofta att man inte riktigt vet om det som presenteras är av bra kvalitét, finns all information, är det rätt information eller saknas någon information, och man upplever att det inte finns någonstans att ta reda på svaren på sådana frågor.
<i>Kan systemet smidigt expandera?</i>	Att expandera är inga problem. Datalagret växer dock och kontinuerlig bortplockning måste göras.	På höjden inga problem, eftersom transaktionerna då växer men inte kuberna dock kan det bli svårt att göra backuper. Svårare på bredden då det blir många dimensioner och stora datamängder.	Kan expandera smidigt genom att lägga till mer hårdvara.	

Tabell 5: Datapresentationsproblem

5 Analys

Här jämför vi de svar fått genom intervjuerna mot de föreskrifter vi har i teorin, om bland annat utbildning för användare, vad verktygen ska klara av, samt användarnas medverkan i hela datalagerprocessen. Vi ger en övergripande bild av förekommande användarproblem.

5.1 Enkät svar

Vi analyserar endast enkätfrågan om **användarproblem**, då det är den som är vår forskningsfråga och syftet är att svara på den. De övriga enkätfrågorna använder vi sedan som de är i slutsatsdragningen.

Användarproblem som framkommit i enkät svaren

A	B	C	D	E	F	G
Långa svarstider	Stora datamängder.	Långa svarstider (löst idag).	Långa svarstider	Integrations-svårigheter.	Långa svarstider	Förståelse för datalager-idén.
Performance-problem.	Data-defini-tions-problem.	Svår-begripligt gränssnitt	Data-kvalitets-problem.		Stora mängder data.	Integra-tions-problem.
			Svåra sökningar		Svåra sökningar	
			Svår-tolkad informa-tion.			

Tabell 6: Användarproblem från enkät svar

Man kan se ett samband mellan långa svarstider och stora datamängder, i så motto att samma företag ofta har problem med bådadera. Ju större mängder data som skall gås igenom vid en sökning, desto längre svarstider ger det. (Detta är i viss mån beroende av hur datalagret är uppbyggt – här kan man påverka datamängd för en sökning.) Andra problem som återkommer är svåra sökningar och performance-problem av olika slag.

Med **integrationsproblem** menas svårigheter med den sammanfogning som måste göras av datalagerlösningens olika delar, då man inte så att säga kan beställa ett komplett datalager. Datalagermiljön måste konstrueras genom att man integrerar de

olika delarna (lagerhanterare, analysverktyg, etc.) med varandra, med hjälp av egen mjukvara.

5.2 Intervjusvar - översikt

Här kategoriserar vi intervjusvaren från tabellen i empiridelen i **ja- och nejsvar**. Detta för att lättare få en översikt över svaren, exempelvis om alla företag svarat samma sak på någon fråga. (Ett par av våra intervjufrågor gick inte av naturliga skäl inte att kategorisera på detta sätt; fråga 13) Vad tror du dessa problem beror på? och fråga 14) Hur tror du dessa problem kan lösas?, gör vi endast en detaljanalys på.)

	Stadium	Findus	Vattenfall	Frontec	Not	Kategori
- Finns inbyggt användarstöd?	Ja	Ja	Ja	Ja		Bristfälligt användarstöd
- Utbildning för att använda verktyg?	Ja	Ja	Ja	Ja		Bristfälligt användarstöd
- Krav och lösningar dokumenterade?	Ja	Ja	Ja	Ja		Bristfälligt användarstöd
- Problem i användandet av datalagret?	Ja	Ja	Ja	Ja		Datautvinningsproblem
- Långa svarstider?	Ja	Ja	Ja	Ja	1)	Datapresentationsproblem
- Bra kvalitet på presenterad information?	Ja	Ja	Ja	Ja		Datapresentationsproblem
- Klarar användaren att göra fullständig sökning själv?	Nej	Nej	Nej	Nej	2)	Bristfälligt användarstöd
- Informationen presenterad i rätt format?	Nej	Nej	Nej	Nej	3)	Datapresentationsproblem
- Verktiget anpassat efter slutanvändaren?	Nej	Nej	Nej	Ja		Medverkansbrister
- Användare med vid inköp av verktyg?	Ja	Ja	Nej	Nej		Medverkansbrister
- Användare deltagit i utvecklingen av datalagret?	Ja	Ja	Nej	Nej		Medverkansbrister
- Expandera smidigt möjligt?	Nej	Ja	Nej	-	4)	Datapresentationsproblem

Tabell 7: Intervjusvar ja/nej

Förklaring till noterna i tabellen:

- 1) Överlag bra svarstider, men alla har erfarenhet av långa svarstider vid stora kuber och komplexare sökfrågor (vilket ibland behövs).
- 2) Med fullständiga sökningar menar vi att användaren själv designar kuberna och avgör vilken information som ska in i dem.
- 3) Frågan syftar på om **verktyget** presenterar i rätt format, inte att man klipper in resultatet i Excel..! (Om man klipper in svaren i Excel, tolkar vi det som nej.)
- 4) Frontecs svar går ej att kategorisera till ja/nej, då vår respondent svarar för flera företag med olika system.

5.3 Intervjusvar - detaljanalys

Här jämför vi – alltefter ”problemkategorierna” - de olika respondenternas svar, dels med varandra och dels med vad som står i litteraturen om den fråga det gäller.

Medverkansbrister:

1) Har du fått vara med och bestämma inköp av verktyg?

Alla fyra företagen har låtit användarna vara med att och påverka inköp av verktyg i olika utsträckning, allt ifrån stor påverkan till ingen påverkan alls. Eftersom Findus anser att det är mycket viktigt att användare finns med så tidigt som möjligt i utvecklingsarbetet, låter man dem påverka inköp av verktyg i form av en användargrupp (representanter för alla användare). Stadium fäster inte samma vikt vid användarmedverkan, men har ändå låtit slutanvändarna välja mellan två verktyg - beslut om vilka verktyg som skulle vara aktuella avgjordes av att ”märket” redan fanns i företaget. Detta kan man även utläsa ur Vattenfalls svar. Vattenfall och Frontecs konsults erfarenheter säger att inga användare fanns med alls vid inköp av verktyg. Enligt Frontecs konsult väljer företag ofta det verktyget som man hört ska vara bra.

Teorin säger att valet av verktyg så mycket som möjligt ska styras av användarens kunskaper om analysarbete och vilka behov av information användaren har (kap. 2.3.2). Teorin säger även att en del verktyg är för svåränvända för beslutsfattaren och att IT-avdelningen ofta väljer verktyg som de själva är förtrogna med, men att detta inte alltid gäller för slutanvändarna (kap. 1.1).

2) Har du deltagit i utvecklingen av datalagret?

Stadiums och Findus användare har fått vara med i utvecklingen av i datalagret redan från början, medan Vattenfalls användare inte varit med alls i utvecklingsarbetet av datalagret. Frontecs konsult menar att det är vanligt att användare inte ingår i utvecklingsprocessen, då projekten ofta är teknikdrivna istället för informationsdrivna.

Teorin säger att det är att rekommendera när man utvecklar ett datalager, att slutanvändaren finns med på seminarier som har med datalagret att göra, och att användarna måste ställa krav. Det är ju slutanvändaren som senare kommer att använda systemet. Man kan säga att systemet måste anpassas till den som ska använda det! Detta eftersom ett datalager utnyttjas effektivast om beslutsfattarna själva kan använda det. Att använda systemet själv klarar man om datastrukturen stämmer med beslutsfattarens egen syn på verksamheten. Användarnas medverkan är mycket viktig i alla faser av design och utveckling av datalagerlösningen (kap. 2.3.1). Om användarna är missnöjda ska projektet anses misslyckat. Missnöjdhet är ofta ett resultat av orimliga förväntningar - något fick användarna att förvänta sig mer än vad de fick, ofta beroende på glapp i kommunikationen mellan IT-avdelningen och användaren (kap. 1.1). En korrekt balans mellan inblandningen av IT-avdelningen och användarskaran är avgörande för slutresultatet. En orsak till att datalagerprojekt inte lyckas är misshandlingen av användarverktyg, man måste ha en speciell metodologi vid framtagningen av dessa (kap. 2.3.1).

3) Är verktyget anpassat efter dig (kundanpassat)?

Stadium och Findus använder sig av fördesignade rapporter efter användarens informationsbehov – själva verktygen är dock inte kundanpassade. Vattenfall Data har inte fördesignade rapporter, men har däremot så kallade datamarts, som innehåller den information användaren behöver. (Själva verktyget är heller inte här individanpassat.) Vår konsult på Frontec säger att det är vanligt med kundanpassning, på så sätt att man anpassar gränssnitten efter de vanligaste sökvägarna.

I **teorin** kan man läsa om Codd's regler för OLAP-verktyg, som säger att OLAP-verktyg skall ge användarna en multidimensionell vy som motsvarar deras bild av företaget och som är "intuitivt analytisk" och lätt att använda (kap. 2.2). Användarna ska delta i utformandet av kravspecifikationen, i datadefinieringen, i valet av slutanvändarverktyg, i utvecklingen av en kundanpassad applikation, i uppföljningen av designdokumentation, i designen av procedurer för att komma åt och manipulera data, i test och validering, i framtagningen av processer för hantering av förändrade krav, samt vid mätningen av användartillfredsställelse med lösningen (kap. 2.3.1). Man kan även läsa om olika faktorer som bör styra valet av klientverktyg en av dessa är anpassning till användaren, man måste fråga sig: Hur generella gränssnitt klarar användaren av? Vilket är behovet av att kunna skapa anpassade dialogformulär? En annan viktig faktor är att ha ett bra, enkelt och lättförståeligt användargränssnitt. Dessa gränssnitt ska vara anpassade till den som ska använda systemet (kap. 2.3.2).

Bristfälligt användarstöd:

4) Klarar du av att göra sökningarna själv?

Stadium och Findus använder sig av fördesignade kuber, eftersom användarna upplever det som svårt att konstruera kuber. Frontecs konsult har också upplevt sådana lösningar

och menar att endast riktigt sofistikerade användare klarar av att konstruera kuber själv. Vattenfall kan inte svara på denna fråga, då man ej gör några fullständiga sökningar.

Teorin säger, att syftet med ett datalager är att individerna själva ska kunna använda det och det är därför mer än vanliga system beroende av individernas acceptans (kap. 2.3.2). Ett datalager utnyttjas effektivast om beslutsfattarna själva kan använda det. Slut användaren ska uppmuntras att använda frågeverktyget självständigt - om det är för invecklat kommer användaren att försöka undvika att använda det (kap. 1.1). De flesta användare av datalagret arbetar vanligen med fördefinierade frågor, rapporter och analyser med eller utan parametrar. Dessa personen ägnar i allmänhet lite tid åt att förstå de möjligheter ett datalager erbjuder (kap. 2.3.2). En orsak till att datalagerprojekt betraktas som misslyckade är att analysarbetet är för komplicerat för användarna (kap. 1.1).

5) Finns inbyggt användarstöd?

Alla företagen vi intervjuat har någon form av användarstöd. Stadium har en form av hjälpdialog. Vattenfall har telefonsupport dit man kan ringa vid problem. Frontec har oftast märkt att man har användarstöd i form av metadata-directory och att detta ofta räcker långt som stöd. Man har oftast även någon typ av telefonsupport.

Teorin säger att utbildning och användarstöd är speciellt viktiga områden. Slut användaren är den som ska trycka på knapparna och helst inte vara i behov av mellanhänder. Det är även viktigt att användaren ska ha tillgång till så kallad metadata, som de själva kan förstå utan att behöva få hjälp av andra eller av någon specialist. En användare som inte har tillgång till metadata kan inte utan hjälp göra analyser. Användare bör även uppmuntras att kommunicera med varandra, via ex intranät eller e-mail för att utbyta erfarenheter. I ett datalager bör det finnas en generell hjälpfunktion för verktyget (kap. 2.3.2).

6) Har du fått utbildning för att använda verktyget?

Alla företag har ordnat någon form av utbildning, men av olika längd. Findus har haft ett utbildningssystem med olika steg beroende på användarens förkunskaper. Stadium har haft en kort internutbildning och därefter tät kontakt med IT-avdelning. Frontecs konsult säger att det ofta förekommer att användare får en kort utbildning på någon timme. Vattenfalla Data har inga direkta kurser som handlar om verktygen, men har däremot kurser i applikationerna som verktygen opererar i.

Teorin talar om tre huvudfaser i datalagerutvecklingen. De två första faserna handlar om hur man tar fram själva databasen, medan den tredje syftar på att göra innehållet i datalagret tillgängligt för slut användarna. Teknisk kunskap tenderar dock att totalt dominera diskussionen om datalager, medan användarproblem inte diskuteras på samma sätt. Följden av detta blir, att användare har svårt att bruka de verktyg som tillhandahålls - de saknar både vana och utbildning att tillämpa analyser och data på rätt sätt och systemet ger inte svar på deras frågeställningar (kap. 2.3.4).

I början av projektet satte man upp särskilda mål för funktion och kapacitet. Dessa kan ha inkluderat datakvalitet, användarutbildning, sättet att leverera data, eller fördefinierade frågor. Om sådana funktioner aldrig realiserades, eller blev uppskjutna till senare tillfällen, kan det betraktas som ett misslyckande (kap.1.1). En sak som man bör tänka på vid valet av klientverktyg är hur lång utbildning som kommer att krävas och hur mycket hjälp kommer användaren att behöva få (kap. 2.3.1).

7) Har man dokumenterat krav och lösningar så att nya användare kan börja använda systemet?

Hos Findus och Vattenfall har man dokumenterat krav och lösningar, och Frontecs konsult har också erfarenhet av att krav och lösningar dokumenteras, så att nya användare kan komma in och börja använda systemet. Dessa dokumentationer finns dock inte till för att nya användare ska komma in och börja använda systemet, utan mer för att de användare som finns idag ska kunna se vad man kan göra. Även Stadium har dokumenterat krav och lösningar, men man anser att dokumentationen är bristfällig och att den heller inte är så viktig.

Teorin säger, att en del av verktyget skall utgöras av dokumentation och att denna skall vara tillgänglig för användaren. Användaren skall även ha tillgång till metadata. När man utvecklar ett datalager bör man redan från början ha stöd och dokumentation (kap. 2.3.2).

Datautvinningsproblem:

8) Upplever du att ni har några problem med att använda datalagret? Vilka problem?

Det stora problemet hos de flesta företag vi intervjuat är allt större mängder data och hur man ska kunna hantera detta utan att det medför problem hos slutanvändarna.

Stadium har t ex problem med **presentation, långa svarstider**, svårt att **söka med många variabler** och överbelastade datorer p g a **stora mängder data**.

Frontec nämner också det här med långa svarstider som en följd av en **allt större mängd data**. Man har även problem med laddtider p g a stora datamängder. Systemens kapacitet uppfattas som för låg för att klara av att ladda in dessa datamängder.

Findus och Frontec berättar om **svårigheter med att få fram data** ur levererande system, men av till synes olika orsaker. Dessa två talar också om svårigheter att riktigt **förstå** det man får ut ur datalagret och om **tveksamheter huruvida den data man söker överhuvudtaget finns** i datalagret. Vattenfall och konsulten på Frontec har båda mött problem med **tilltron** till den data man får - är den data man får ut rätt och riktig?

Mer än en av våra respondenter nämner svagheter i själva systemet - det har **för låg kapacitet** för att klara att hantera större mängder data. Vi har även under intervjuerna fått fram, att mer kunniga användare tycker att man gör det för lätt för sig genom att bara skylla på verktygen - det är oftast inte själva verktygen som problemen beror på.

Teorin säger att de vanligaste problemen är just stora mängder data, som i sin tur orsakar långa svarstider, långa laddtider och komplexa sökningar (kap. 2.3.3-2.3.4).

Konsistens i ”performance” – antalet dimensioner, aggregationsnivåer och storleken på databasen ska inte påverka verktygets ”performance” nämnvärt; rapportfigurer och uppställningar ska inte förändras till det sämre av att dessa blir fler; obegränsade dimensions- och aggregationsnivåer – man ska kunna ha många nivåer och dimensioner av data (Codds regler, kap. 2.2!). Man bör fråga sig, hur stora datavolymer som kommer att behöva genomsökas för att man skall kunna utföra aktuella sammanställningar (kap.2.3.1).

Utan analytisk kunskap kan så kallade ad-hoc-verktyg producera osäkra resultat. Frågor som dessa är inte ovanliga: Jag är inte säker på vad jag frågade efter! Hur många objekt kommer jag att få som svar? Har jag ställt en motsägelsefull fråga? Är frågan tillräckligt specificerad? eller; Hur kan jag kontrollera att resultatet stämmer? (kap. 2.3.2).

Viss data som man behöver finns inte i de existerade källsystemen (kap.2).

9) Vad tror du dessa problem beror på?

Respondenterna har angett en del likartade problem, men orsakerna till dessa tycks olika. Stadium menar att problemen till stor del beror på för stora mängder data för att datorerna ska klara av att processa dem. Många av Findus problem beror på att de är ett internationellt företag och det finns svårigheter i att samla data från alla länders datakällor till en gemensam databank. En orsak till problemen är att systemen i andra länder har dålig kapacitet, man har svårigheter i att förstå data som kommer från ett annat land och att kvalitén ibland kan vara dålig och åtkomligheten kan vara svår. Vattenfall tror att bristande intresse från användare och försämrade prestanda beror på att datalagret modellerats felaktigt. Frontecs konsult menar att politik inom organisationerna kan vara roten till många problem, såsom exempelvis dataäggande. Vissa problem kan bero på att man är ovan vid systemet och att det ibland även kan vara svårt att börja med något nytt. En annan orsak kan vara att datalager projektet började som en teknikdriven "fluga", d v s **inte** styrdes av krav från beslutsfattarna.

10) Hur tror du dessa problem kan lösas?

Både Findus och Frontec tar upp, att problem med långa svarstider kan lösas genom att ”trimma” OLAP-servern (för att snabba upp svarstiderna) eller genom att ta reda på användarens vanligaste sökvägar och snabba upp dem.

Findus har tidigare haft problem med att användare dels inte kan göra sökningar själva och dels inte förstår det de får ut datalagret. Detta har man nu löst genom ett **webbverktyg**, samt **fördesign** av sökningar. Man har även tidigare haft problem med för stora kuber, vilket i sin tur gett långa uppdateringstider, svårigheter med att göra backuper och att kuberna tagit upp mycket utrymme. Detta problem har man löst genom att skapa mindre kuber och att använda sig av delkuber. En liknande lösning har

presenterats av Stadium. Problemet med kuber som tagit för stort utrymme har dessa löst genom att skapa så kallade **virtuella kuber** – man behöver därigenom inte processa två hela kuber, utan bara vissa dimensioner av dem. Problem med långa svarstider har lösts tillfälligt genom att köra analyserna direkt på servern, och långsiktigt genom att skapa **index** för vissa frågor som man vill ställa återkommande, d v s **snabba upp sökvägarna**. Att snabba upp sökvägarna är även för Frontecs konsult en lösning på de långa svarstiderna. Hon tror även att problemet med dataäggande kan lösas genom att man använder sig av så kallad **datasteward** (som handhåller data och avgör vem som skall ha tillgång till vad). Frontec menar även att man - för att underlätta tolkningen - kan skapa **enkla** användargränssnitt med **stor flexibilitet**. Vattenfall tror att tillförlitligheten till den data man får ut kan förbättras genom **kvalitetskontroller**. Prestandan kan optimeras genom **ändrade lagringsformer**.

Datapresentationsproblem:

11) Får du svaren i rimlig tid?

Under intervjuerna har det framkommit att alla intervju företag har eller har haft problem med långa svarstider.

Stadium och Findus har idag överlag bra svarstider. Dock tycker Stadium, och även Vattenfall, att svarstiderna blir långa då större datamängder skall genomsökas. Frontec tycker det är vanligt med långa svarstider, men att det ofta beror på att användaren ställt frågan felaktigt.

Teorin säger att datalagerprojekt ofta avblåses på grund av oacceptabel ”performance”. ”Performance” syftar både på svarstid för frågor och på tid för inladdning av data till datalagret. Användaren måste förstå vad man kan förvänta sig, men samtidigt finns det gränser för vad man skall acceptera! (kap. 1.1).

Ju större ett datalager blir, desto längre brukar svarstiderna bli.

12) Är informationen presenterad i det format du vill ha det i?

Alla tillfrågade har svarat att man använder sig av Excel för att lättare kunna tolka, analysera och förstå resultatet. De grafer som verktyget presenterar resultatet i uppfattas oftast som svåra, eller så känner man bara att man kan Excel bättre.

Teorin säger att en viktig faktor är att ha ett bra, enkelt och lättförståeligt användargränssnitt. Dessa gränssnitt ska vara anpassade till användaren (kap. 2.3.2). En del verktyg är för svåränvända för beslutsfattaren. IT-avdelningen väljer ofta verktyg som de själva är förtrogna med, men detta gäller inte för slutanvändarna! Om verktyget är för invecklat, kommer slutanvändaren att försöka undvika att använda det (kap. 1.1). Detta stjälper själva syftet med ett datalager, att slutanvändaren ska uppmuntras att använda frågeverktyget självständigt. Data i ett datalager ska i första hand ordnas på ett sätt att den blir förståelig för slutanvändaren. Idag används olika

gränssnitt för att presentera information, det finns färdiga rapporter, färdiga formulär, färdiga frågor och utnyttjande av klientverktyg själv direkt mot en generell struktur (kap. 2.3.4).

13) Är kvalitén på det som presenteras bra?

Samtliga respondenter anser att kvaliteten på det de får presenterat för sig är bra. Det kan dock, säger Frontecs konsult, vara svårt att veta om kvaliteten är bra eller inte – t ex kan data saknas - eller finns inte ens i datalagret! - utan att man vet om det.

Teorin säger att det är viktigt att fråga sig vilken kvalitét data har. Idag finns olika funktioner för att garantera att data är korrekt och tillgänglig. Det är viktigt att skapa ett förtroende för den informationen som hämtas ur datalagret - ofta är användarna skeptiska till den informationen de får ut, och detta med rätta! Ansvar för den eller de personer som ska ta hand om datakvaliteten ska definieras och användarna ska löpande informeras om kvalitén (kap. 2.3.2).

14) Kan systemet smidigt expandera?

Alla tillfrågade företag har svarat att det inte är några problem att expandera systemet - i vart fall inte på höjden, eftersom transaktionerna då växer, men inte kuberna. Vattenfall tycker även att det går bra att expandera på bredden, genom att lägga till mer hårdvara. De andra företagen anser dock att det är svårare att expandera på bredden, då det blir många dimensioner och stora datamängder. Stadium tycker dock inte att den lösning man har nu är så bra - man får hela tiden plocka bort data i ena änden av lagret för att kunna lägga till i den andra - man kan alltså säga, att det **inte** går så smidigt att expandera.

Teorin menar (kap. 2.3.3-2.3.4), att den stora utmaningen inför framtiden är att kunna låta datalagret växa (på grund av att organisationer använder sig av allt större datamängder) utan att detta inverkar alltför mycket på svarstiderna. Redan idag anser många slutanvändare att svarstiderna för sökningar är för långa. Det är samtidigt en fråga om vad man kan förvänta sig; användare utan teknisk kunskap måste bli upplysta om vad det är rimligt att vänta sig för svarstid på t ex en sökning som kräver genomgång av hela datalagret. Den stora saken i informationsprocessen är hur man ska kunna processa större och större databaser utan att få långa svarstider.

Ett datalager kan ta stort diskutrymme. Om den faktiska datan är i flera dimensioner är kombinationer av tabeller och index till den faktiska datan, kan ta upp mer utrymme än rådatan (kap 2).

Datalager innebär ofta stora datamängder. Även om parallell teknik och snabba datorer underlättar förbättrar aggregering, partitionering, kodade attribut och indexering prestanda. Det gäller att ta ställning till hur länge data ska lagras i datalagret (kap. 2.3.3).

6 Diskussion och slutsatser

Syftet med vårt arbete har varit, att ge en samlad bild av förekommande svårigheter för slutanvändare av datalagret. Den forskningsfråga som skulle besvara för att uppnå syftet var: **Vilka problem kan uppstå för slutanvändare när de tillämpar datalagret och vad beror dessa problem på?**

6.1 Litteratur

Vi tycker att vi fått svar på vår forskningsfråga genom undersökningen och därmed uppnått vårt syfte. En slutsats vi kunnat dra från den litteraturstudie vi gjorde innan undersökningen började är, att det finns många fördelar med datalager, men också en del problem, som det kan ta år att komma tillrätta med. Det tycks också finnas ett medvetande om vilka användarproblem som kan uppstå och hur man skall gå tillväga redan i utvecklingsarbetet för att undvika dem (litteraturen är till stor del av ”förebyggande” karaktär, med punkter för hur man skall arbeta, o s v).

Den litteratur vi läst har varit av blandade årgångar - den äldsta från mitten av 90-talet - och vi har noterat att den nyaste litteraturen (från 2000-talet) talar om samma problem; man får intrycket att det inte hänt mycket på flera år vad gäller många av dem. En annan iakttagelse vi gjort är, att det endast finns enstaka böcker om datalager skrivna på svenska – den mesta litteraturen är skriven av amerikanska författare. Det tycks alltså som att datalager fortfarande är ett ganska obeprövat område i Sverige, trots att allt fler använder det.

Det problem som tydligast framkommer i litteraturen är performance-problem av olika slag, däribland långa svarstider för sökningar i större datalager (kap. 1.1, 2.3.3 och 2.3.4), och det finns ännu inte någon ”mall” för att komma tillrätta med det – det betraktas som en utmaning.

6.2 Fallstudie

När vi tittar på de enkätsvar vi fått, ser vi, att företagen har likartade svårigheter för slutanvändare trots att man använt sig av datalager så olika länge (från endast ett par år upp till sex år). Detta kan bero på en rad omständigheter, som t ex att problemen är så pass svårlösta att det tar flera år att ”bygga bort” dem, eller att företaget inte satsar tillräckligt med tid och pengar på att förbättra villkoren för slutanvändare av datalagret. Omvänt kan det vara så, att slutanvändarna till en början är så nöjda med det nya systemet (som underlättar så mycket) att man inte lyfter fram problemen för dem som eventuellt kan åtgärda dem!

Vår undersökning har vidare bekräftat det litteraturen tar upp; att datalageranvändningen är behäftad med en rad problem (men att företagen överlag tycker att det är en bra lösning). Det mest framträdande problemet som framkommit i under-

sökningen är långa svarstider vid sökningar (se Empiri och Analys!). Detta kan i sin tur vara orsakat av stora mängder data man ej klarar att hantera på ett smidigt sätt (är stora datamängder felstrukturerade, måste onödigt många gånger gå igenom vid sökning och svarstiden växer), eller att man helt enkelt gjort sökningen felaktigt. Nämnas kan, att det bara var ett av våra intervjuföretag som hade en uttalad strategi för hantering av stora datamängder. Övriga var tvungna att ideligen plocka bort "gammal" data i ena änden av datalagret, för att kunna lägga till nyare i den andra. Den sistnämnda orsaken till långa svarstider tycks dock vara den vanligaste. Vid intervjuerna har det kommit fram, att användare önskar sig möjligheter att i förväg kunna bedöma hur länge en sökning skall komma att ta, men att detta inte är möjligt i dagsläget. (Det skulle, på så sätt, vara mycket lättare att planera sitt arbete med datalagret.)

6.3 Undersökningsspecifika slutsatser

Vi har också sett, att analysverktygen är så komplicerade att använda för den mindre skolade slutanvändaren, att många företag har lyft bort det ansvaret från användarna genom att låta tekniker fördesigna rapporter och presentationer. På grund av knapphändig utbildning i verktyget, eller brister hos verktyget, har användarna svårigheter med tolkningen av den information verktyget presenterar för dem. Följden blir ofta, att användaren "klipper in" sökresultatet i Excel, som man är van vid och tycker har ett lättare sätt att redovisa data. Trots att samtliga våra respondenter uppgav denna form av användning vågar vi inte anta att det gäller för alla svenska företag (se Empiri, samt Analys : Intervjusvar – översikt!).

Det är inte vanligt att användarna funnits med vare sig i utvecklingsarbetet med datalagret, eller vid inköp av analysverktyg. Det förekommer att datatekniker på företagen rentav tycker att det är besvärligt och tidskrävande att inkludera användare. Valet av verktyg styrdes vanligen av produkttillverkare som redan fanns på företaget av andra anledningar – man bestämde sig direkt för ett datalagerverktyg från dem av enkelhetsskäl. Det är likväl få företag, vid vilka krav/databehov hos slutanvändarna var det som satte igång själva datalagerprojektet. Istället var det enstaka, tekniskt intresserade anställda som drev idén om att man skulle upprätta ett datalager. Dessa hade också redan från början förslag på vilket företag/vilket märke på datalager man skulle välja – det var det namn som så att säga var i ropet och skulle vara bra, som blev aktuellt (se Analys : Detaljerad analys!).

6.4 Allmänna slutsatser

Med denna rapport vill vi ge följande **kunskapsbidrag: större medvetande** om de problem som lätt uppstår för slutanvändare av datalager. (Om man känner till vilka problem som kan uppstå och försöker bygga bort dem från början blir datalagerprojektet inriktat mot att få lösningen så bra som möjligt för slutanvändaren. Detta arbete ser - till skillnad från de flesta andra undersökningar - på **svårigheter** med datalager, istället för enbart vinsterna.)

Följande slutsatser har vi dragit, som torde gälla för de flesta företag:

- **Problem ur användarperspektiv finns** (se Empiri!) - dessa kan dock vara olika och av större eller mindre proportion. Samma problem har återkommit i enkätsvar och vid intervjuerna. Vi tror att om vi skulle ha frågat fler företag skulle vi ha sett samma saker.
- **Långa svarstider vid sökningar** (se Empiri och Analys!) – litteraturen och talar om detta problem och samtliga våra respondenter upplever att svarstiden växer med kubstorlek och frågans komplexitet. Vi drar här slutsatsen att det saknas flexibilitet i sökningar – så snart man vill gå utanför det förutbestämda informationsbehovet och utöka sökningen, får man långa svarstider.
- Någon form av **användarstöd finns** och ett telefonnummer att ringa för teknisk support (se Empiri och Analys!). Att detta är en skillnad mot litteraturen kan bero på att verktygen idag är modernare än när litteraturen skrevs – respondenterna har i sina verktyg färdiginbyggda stöd och eftersom verktygen de använder är vanliga tror vi att även andra företag har detta stöd.
- **Utbildning i verktygen** är praxis idag, och anses som tillräcklig (se Empiri, Detaljanalys) – fastän litteraturen talar om att utbildningen i verktygen är otillräcklig tycker våra respondenter (med helt olika bakgrund och erfarenhet) att utbildningen räcker.

6.5 Summering av egna reflektioner

Litteraturen talar om vikten av att användare tas med redan från början, men att det är mindre vanligt att företagen tänker på detta. Man kan dock, menar artikelförfattare, föregå/undvika många problem som förr eller senare skulle uppstå, genom att ta med användare från början (Söderström (1997), Bischoff & Alexander (1999)).

Om man redan från början intresserar sig för användares informationsbehov och synpunkter, har man också en hel del vunnet när det gäller att skola in användarna på det färdiga systemet. Vi kan notera att - i kombination med att få slutanvändare inkluderats i datalagerprojektet – alla våra företag endast gett sina datalageranvändare en kortare grundutbildning (på högst ett par timmar), därefter har de själva fått sitta och prova sig fram, samt tillfråga varandra. Detta är kanske anledningen till att – fastän användarna själva anser att deras utbildning varit tillräcklig – de ändå har svårt att göra en fullständig sökning själv. Detta har istället ”smugits bort” genom att tekniker på företaget fått förmedlat vilka informationsbehov som finns och fördesigna rapporterna åt användaren (se Empiri!).

När slutanvändare blir tvungna att i efterhand förmedla sina informationsbehov till tekniker (som kan IT, men ej har tillräcklig kunskap om företaget), uppstår lätt det så kallade informationsgapet (se avsnittet Teori). Detta kan medföra, att de som är experter på företaget och dess behov inte riktigt får den lösning de tänkt sig; datas betydelse för företaget kan missförstås av teknikern, det kan saknas tillgång till data som användaren egentligen efterfrågat, o s v. Detta tror vi kan ha drabbat våra företag i

större utsträckning än de själva är medvetna om, helt enkelt på grund av den mänskliga faktorn vid informationsöverföring.

Det har nämnts tveksamheter om huruvida man kan förlita sig fullt ut på den data man får fram vid sökning. Finns viss data över huvud taget i datalagret, är den data som kommer upp rätt och riktig, eller saknas det någon data i det resultat man får vid en sökning? Det har framförts önskemål om att detta skall gå att kolla upp – det gör det inte i dagsläget för den vanlige användaren.

På frågan om de data som presenteras för användaren är av bra kvalitet, svarade samtliga respondenter ja. Vi kan dock notera, att alla samtidigt använder sig av Excel (som är mer bekant och anses lättare att förstå) för att presentera data. De uttalar sig därmed om hur kvalitén är på en presentation i Excel – inte i verktygets eget gränssnitt (se Empiri!). Det kan alltså vara som vi tror, att det finns brister i kvalitén på data som presenteras genom verktyget, som företaget inte ser.

För den användare som stöter på svårigheter, finns för våra respondenter oftast någon form av inbyggt användarstöd. Det har dock visat sig, att många slutanvändare hellre direkt ringer till en tekniker, som oftast finns i närheten. Vissa användare ringer också direkt till leverantören av verktyget. På så sätt får de snabbare svar, än om de ringer ett hjälpsnummer eller dylikt. (Våra respondenter känner också – till skillnad från vad som sägs i litteraturen – att de har god kontakt med tekniskt kunniga och kan be om hjälp när problem uppstår – se Empiri!)

Följande **skillnader** kunde vi se mellan litteraturen och undersökningsresultatet:

- **Användarstödet** på våra företag är en positiv skillnad mot vad som står i litteraturen – där är det oftast långt till hjälpen och det finns inte heller inbyggt användarstöd i verktyget (kap. 2, 2.3.2).
- Användarna i undersökningen är **nöjda med den tid som avsatts för utbildning** i verktygen - litteraturen talar om att utbildningstiden ofta upplevs som för kort (kap. 1.1).

Något som det varit svårt att dra några generella slutsatser om, är **orsaken** till de problem företagen har. Våra intervjuobjekt har nämnt likartade problem för slutanvändare, men dessa har olika orsaker. Då orsakerna är olika, blir också den tilltänkta lösningen på problemen olika. Detta är kanske en anledning till att dessa problem fortfarande finns – man vet att man har problem, men då det inte finns någon gängse lösning på dem, förmår man inte riktigt komma tillrätta med dem (se Empiri!).

6.6 Förslag till fortsatt forskning

Vi har undersökt problem som kan uppstå för slutanvändare av datalager. Bland de problem som framkommit, finns sådana som egentligen har sin rot i **inladdningen av data från källsystemen** (se kap. 2!) - t ex performance-problem och bristande tilltro till data - samt i tvätten av rådatan innan den läggs in i datalagret. Vi skulle föreslå en annan grupp att titta på denna del av datalagermiljön – alltså **”front-end”-komponenten** (se bild i Teori). Den är behäftad med minst lika mycket problem som den del vi undersökt. Ett annat förslag är att fokusera på en av våra intervjufrågor, som inte tillhörde huvudfrågorna, men som är väl så intressant: **Hur kan dessa problem lösas?** Vi har i vissa fall endast frågat slutanvändaren hur denne tror att problemen kan lösas – vidare forskning kan istället gå till djupet med denna fråga och höra med de som verkligen har kunskapen.

Avslutningsvis vill vi trycka på, att de problem för slutanvändare som vi undersökt är små i förhållande till de enorma förtjänster ett datalager har för företaget. Alla våra respondenter har betonat detta. Dock skall problemen inte förbises – ju snabbare man upptäcker och tar tag i små problem, desto bättre och snabbare kan de lösas och desto mindre är risken att de blir stora och svåra att komma tillrätta med.

7 Litteraturförteckning

Böcker:

Bischoff, J. Alexander, T. *Data Warehouse – Practical advice from the experts*, Prentice Hall, Inc. Upper saddle River, New Jersey. 1997. ISBN: 0-13-577370-9.

Connolly, T. Begg, C. *Database systems*, Addison Wesley Longman Limited, Essex. 1999. ISBN: 0-201-34287-1.

Söderström, P. *Data Warehouse – Datalager*, DF Förlags AB, Studentlitteratur, Lund. 1997. ISBN: 91-88862-04-6.

Thurén, T. *Vetenskapsteori för nybörjare*, Liber AB, Stockholm. 2002. ISBN: 91-47-14807-7.

Yin, R. *Case study research: Design and Methods*, Sage Publications, Inc. 1994. ISBN: 0-8039-5662-2.

Rapporter:

Hortell, P/Soini, P. *Datavaruhus*. Luleå: Luleå Tekniska Universitet, Institutionen för industriell ekonomi och samhällsvetenskap, avdelningen för data- och systemvetenskap, 1999:35.

Holmgård, F/Fazeli, F. Luleå: Luleå Tekniska Universitet, Institutionen för industriell ekonomi och samhällsvetenskap, avdelningen för data- och systemvetenskap, 2003:.

Webbdokument:

Alur, N. White, C. *Data warehousing :Directory-Driven Information Delivery*. Data Base Associates International. www.dbaint.com , (July) 1996.

Business Navigator, *Datalager: Enklare och billigare*, www.busnav.se, 1998.

Next Action Technology (sales@nextaction.co.uk), *Filling the Fact Gap*, www.AnswerSets.com , 1997.

Seiner, R S, *Data Administration Newsletter*, www.tdan.com , (February) 2001.

Bilaga A - Enkätfrågor

1. Använder ni datalager (Data Warehouse)?

Ja []
Nej []

Kommentarer.....
.....

2. Hur länge har ni använt er av det?

.....år

3. Har ni upptäckt några problem (ex svårt att få fram data, svårt att göra sökningar eller långa svarstider) vad gäller användandet av verktygen?

Inga problem alls []
Vissa problem []
Mycket problem []

Kommentarer.....
.....

4. Hanterar och underhåller ni själva datalagret eller finns personal från utvecklingsföretag hos er för detta?

Vi sköter det helt själva []
Vi hyr in personal för detta []

Kommentarer.....
.....

5. Om vi vill veta mer hur kan vi nå er/dig (ange ex. e-mail adress, telefon nr)?

.....
.....

6. Är det något övrigt som du vill tillägga?

.....
.....

Tack för er medverkan, vi hör av oss om vidare kontakt!

Bilaga B - Intervjufrågor

- Hur fungerar er datalager idag?
- Vilken information behöver du för att fatta beslut? Vad av detta får du ut idag?
- Finns det information som du skulle vilja få ut ur datalagret men inte får idag?
- Upplever du själv att du har några speciella problem med att använda datalagret? Vilka problem? Vad tror du att dessa problem beror på?
- Har ni tidigare haft några problem med informationsutvinningen ur datalagret?
- Vad har du fått för utbildning för att använda verktygen? Var utbildningen tillräcklig?
- Har du deltagit i utvecklingen av datalagret?
- Är verktyget anpassat (kundanpassat) efter dig?
- Har du fått vara med och bestämma inköp av verktyg?
- Är kvaliteten på det som presenteras god? Är det lätt att göra rapporter?
- Är informationen presenterat i det format du vill ha det i ?
- Får du svaren inom rimlig tid?
- Klarar du av att göra sökningarna själv? Är det lätt att göra sökningar?
- Skulle du önska att det fanns inbyggt användarstöd?
- Kan systemet smidigt expandera?
- Hur är relationen mellan er tekniker och slutanvändare?
- Har man dokumenterat krav/lösningar och metadata så att nya användare kan komma in och börja använda systemet?