



Handelshögskolan
VID GÖTEBORGS UNIVERSITET

Big Data

- från tekniskt problem till ekonomisk möjlighet?

Magisteruppsats i företagsekonomi
Ekonomistyrning
Vårterminen 2012
Handledare: Urban Ask
Författare: Henrik Magnusson 84
Johan Viggedal 84

Sammanfattning

Examensarbete i företagsekonomi, Handelshögskolan vid Göteborgs universitet, Ekonomistyrning, Magisteruppsats, VT 2012

Författare: Henrik Magnusson, Johan Viggedal

Handledare: Urban Ask

Titel: Big Data – från tekniskt problem till ekonomisk möjlighet?

Bakgrund och problem: I takt med att allt mer information genereras har en problematik uppkommit kring hur man skall hantera denna i organisationernas IT-infrastruktur. Utbudssidan hävdar att deras nya lösningar kommer att revolutionera möjligheterna kring analys av alla olika former av data. Tillgång bättre information och förbättrade analysmöjligheter kan få stora konsekvenser för ekonomistyrningen förutsatt att det används på rätt sätt.

Syfte: Vårt syfte är att identifiera och beskriva innebörden av begreppet Big Data samt hur de teknologier och verktyg som finns kan stödja ekonomistyrning. För att sätta området i ett större sammanhang har vi studerat begreppet utifrån tre aktörers perspektiv.

Metod: En kvalitativ undersökning med intervjuer av respondenter med kunskap inom området har genomförts. Det empiriska materialet har analyserats utifrån en referensram bestående av fakta om området Big Data och teori om ekonomistyrning.

Resultat och slutsatser: Big Data kan sammanfattas som den utmaning som uppstått kring hantering av datamängder inom volym på data, variation av datatyper och hastighet på datagenerering. Det finns en samstämmighet hos leverantörer, konsulter och användare av stora mängder data att detta är en existerande problematik vars lösningar innebär möjligheter. Förbättrad möjlighet till analys av stora mängder strukturerad och ostrukturerad data kan stödja ekonomistyrningens uppföljning, mätning och prognostisering. Även beslutsunderlag och uppslag kring strategiska beslut och effektivisering av processer genom automatisering kan leda till en effektivare ekonomistyrning. Analysarbetet kan dessutom påverkas när man går från att analysera det man på förhand definierat som viktigt till att analysera sådant man inte visste kunde vara viktigt.

Förslag till fortsatt forskning:

- Göra en liknande studie om några år för att se om, eller hur, den breda massan tagit till sig tekniken och dess möjligheter. Vilka företag och branscher använder teknologin? Har teknologin växt in i det större området BI eller betraktas de fortfarande som separata?
- Utföra djupare fallstudier i organisationer som använder sig av den här typen av lösningar för att studera hur de arbetar praktiskt och hur deras ekonomistyrning påverkas av möjligheterna till nya typer av analyser.
- Göra en studie i syfte att beskriva hur analysarbetet i organisationer utvecklats med hjälp av IT-stöd i ett större perspektiv, till exempel i relation till Skirletz modell över operativt och strategiskt arbete med BI.

Innehållsförteckning

1. Inledning	1
1.1 Problemformulering	4
1.2 Frågeställning	5
1.3 Syfte	5
2. Metod	6
2.1 Val av metod	6
2.2 Ekonomistyrningsinnovationer	7
2.3 Datainsamling	7
2.2.1 Primärdata	7
2.2.2 Sekundärdata	7
2.2.3 Intervjuprocessen	8
2.2.4 Urval	8
2.3 Giltighetsanspråk	9
3. Referensram	11
3.1 Big Data	11
3.1.1 Relation till Business Intelligence	13
3.1.2 Strategisk och operativ BI	14
3.1.3 Var skapas informationen	15
3.1.4 Design och paketering	17
3.1.5 Teknik	19
3.2 Ekonomistyrning	21
3.2.1 Management control package	21
4. Empiri	27
Case 1: Vestas Wind Systems	27
Case 2: Station Casinos	27
Case 3: AFA Försäkring	28
Case 4: King	29
4.1 Intervjuer	30
4.1.1 Introduktion	30
4.1.2 Begreppet Big Data	31
4.1.3 Användning och designelement	33
4.1.4 Retorik och behov	37
4.1.5 Sammanställning av intervjuer	41
5. Analys	42
5.1 Begreppet Big Data	42
5.2 Relation till BI	43
5.3 Användning	44
5.3.1 Teknik	44
5.3.2 Praktik utifrån casen	45
5.4 Ekonomistyrning	47
5.4.1 Big Data lösningar som ekonomistyrningsinnovationer	49
5.4.1.1 Designelement	51
5.4.1.2 Retorik	52

6. Slutsatser	54
6.1 Vad innebär Big Data enligt leverantörer, konsulter och användare?	54
6.2 Hur kan Big Data stödja ekonomistyrningsarbetet?	55
7. Avslutande diskussion	57
8. Förslag på fortsatt forskning.....	58
9. Referenslista	59

1. Inledning

De senaste decennierna har användning av IT ökat kraftigt och spritt sig i alla delar av samhället. Alla företag använder IT i någon form och det finns överallt från industrimaskinernas sensorer till apparater i våra hem. Dessutom har användningen av internet växt med nya tjänster och koncept såsom web 2.0, detta har lett till mer användargenerad information (Tredinnick, 2006). Microblogg-siten Twitter rapporterade under 2011 att man nått en nivå av över 200 miljoner användargenerade inlägg om dagen (Twitter, 2011) samtidigt innehåller den sociala nätverkssidan Facebook i dagsläget över 50 miljarder bilder och generar över 130 terabyte i loggade användarhändelser per dygn (Johnson, 2010). Vid sidan av detta innehåller en stor del av dagens nya produkter, fordon och mobiltelefoner så kallat "smarta" element och loggar mängder med information. Datajätten IBM talar om hur 90 % av den data vi har idag är genererad enbart under de senaste två åren (IBM, 2012). Vi bombarderas varje dag med information vare sig vi vill eller inte och enligt the Economist (2010) utsätts eller hanterar den genomsnittliga Amerikanen för ungefär 34 gigabyte data varje dag i sitt vardagsliv.

Vare sig vi använder oss av mobiltelefoner, handlar med betalkort eller flyger är vi med och bidrar till att mängder av data genereras hos operatörer, tillverkare, tjänste- och produktleverantörer. Världens största detaljhandelskedja, Wal-Mart registrerar en miljon kundtransaktioner varje timme. En flygning över atlanten med ett fyrmotorigt jetplan kan generera ungefär 640 terabyte data. Ser vi detta ur ett perspektiv med de över 25000 flygningar som sker i världen varje dag inser vi fort vilka otroliga kvantiteter det kan röra sig om bara inom en bransch. (Rogers, 2011).

Denna utveckling har enligt det amerikanska konsultföretaget McKinsey (2011) lett till vad man kallar för en dataexplosion där företag har tillgång till enorma mängder data som kan vara en viktig källa till konkurrensfördelar och strategiskt övertag. Den som har tillgång till detaljerad information om sin egen verksamhet och omvärld kan givetvis dra stora fördelar av det om man kan hantera den på rätt sätt. McKinsey uppskattar att den globala datavolymen kommer att växa 44 gånger om mellan 2009 och 2020. En annan bidragande orsak till detta vid sidan av spridningen av teknik är den allt billigare lagringen. Kostnaderna för lagring per gigabyte har sjunkit kraftigt

de senaste decennierna. Dollarpriset per gigabyte har sjunkit från cirka 200,000 dollar år 1980 till dagens priser under tio cent (Komorowski, 2009).

Att analysera informationsmängder i syfte att stödja organisationen är inget nytt och begreppet Business Intelligence (BI), som Gartner definierar enligt nedan, har varit på branschens agenda i snart ett decennium med ett stort genomslag.

“...an umbrella term that spans the people, processes and applications/tools to organize information, enable access to it and analyze it to improve decisions and manage performance.” (Gartner, 2012)

BI-plattformar gör det möjligt för användarna att bygga applikationer vilka hanterar den data som beskriver organisationens interna och externa verklighet baserat på den data som finns tillgänglig.

Men dagens datahantering i kombination med de kraftigt ökande informationsmängderna har lett till problem. Aktörer i branschen har en uppfattning att traditionella databaser och verktyg för statistisk analys inte räcker till (Russom, 2011) för att plocka ut och sortera relevant information med en tillräcklig hastighet i det fall det innefattar stora datamängder och/eller så kallad ostrukturerad data som inte kan sorteras in i en relationsdatabas. Kring detta har termen Big Data, vilken myntades av analytikerfirman Meta group 2001, vuxit fram. Big Data är ett brett begrepp vilket har en rad olika definitioner men det syftar generellt till datamängder som är för stora eller komplexa för att hantera och analysera inom en realistisk tidsrymd med existerande teknik (Prentice, 2011).

Dessa infrastrukturella problem i kombination med affärsmöjligheterna som den lagrade informationen utgör innebär givetvis att många intressenter fått upp ögonen för området på både utvecklings-, leverantörs- och konsultsidan. Ledande teknikföretag har drivit sina egna projekt och leverantörer av IT-lösningar arbetar idag med att utveckla, samla och paketera tekniker som möjliggör användande av den ständigt ökade informationsmassan.

Följer man statistiken för användningen av sökordet "Big Data" från Google insights i figuren nedan ser man att intresset har ökat explosionsartat det senaste året. Detta talar för att ämnet är väl värt att undersöka.



Figur 1. Google trender över sökningar på begreppet "Big Data" från mars 2010 till mars 2012

Källa: (Google)

Analytikerfirman Gartner talar om Big Data som ett av tre huvudsakliga utvecklingsområden för BI de kommande åren (Bitterer, 2011) tillsammans med bredare adoption och utveckling inom mätning/beslutsfattande. Det Big Data-relaterade ramverket Hadoop nämns dessutom som en av de tio teknologitrender som bedöms att få stor betydelse för branschen under 2012 (Casonato, et al. 2011).

Innovation Area	Technologies	Rationale
Consumerization	Search, mobile, visualization, data discovery	To drive adoption
Decision support	Collaborative decision making, predictive analytics	To extend the scope of BI beyond measurement and classification, and to apply information directly to decision making
Non-traditional data, "big data"	Content, text analytics, semantics, in-memory analytics, in-memory DBMSs, columnar DBMSs	To expand the scope of BI to new data types and sources
BI = business intelligence; DBMS = database management system		

Figur 2. Innovationsområden för Business Intelligence enligt Gartner

Källa: (Bitterer, 2011)

1.1 Problemformulering

Det är tänkbart att det ökade informationsutbudet innebär tillgång till flera typer av information vilket kan ge nya förutsättningar för ekonomistyrning. Det är sedan länge fastslagit att bra information som hanteras på rätt sätt innebär fördelar för företag (Porter & Millar, 1985). Samtidigt har det länge talats om vikten att inkludera nyckeltal och mål vilka inte har sitt ursprung i det rent finansiella i syfte att skapa en mer balanserad styrning (Kaplan, 1983). Utbudssidan har ett stort inflytande i vilken väg utvecklingen av tekniska lösningar skall ta och de menar att Big Data-lösningar är ett av de mest expansiva områdena inom branschen (Casonato, et al. 2011).

Information är i större utsträckning omtalat som en källa till strategiska konkurrensfördelar och utbudssidan argumenterar därför med kraftig retorik att det är mycket viktigt att ta tillväxten i data på allvar (Prentice, 2011). Enligt Gartner (Prentice, 2011) riskerar företag som inte adopterar den nya tekniken för hantering och analys av stora och ostrukturerade datamängder att bli ifrånsprungna av sina konkurrenter. All tillgänglig information kan skapa möjligheter till exempelvis bättre styrning eller effektivisering, vågar man riskera att gå miste om detta?

Carr (2003) beskriver att företag kan ha fördelar gentemot sina konkurrenter om de har bra insikt i hur nya teknologier ska användas. Samtidigt menar han att man ska akta sig för att tro att dessa fördelar kommer att bestå. Fördelar gentemot konkurrenter blir sällan långvariga, i alla fall när det gäller tekniker som inte kan skyddas med till exempel patent eller äganderätt. IT-lösningar tenderar enligt Carr att

vara en sådan teknologi. Än så länge har marknads aktörer inte samlats kring ett tydligt koncept kring vad Big Data är och vad det innebär för företag och dess styrning. Vi vill ta reda på mer om detta nya och snabbväxande område, därför har vi formulerat nedanstående frågeställning för denna uppsats.

1.2 Frågeställning

Vad innebär Big Data enligt leverantörer, konsulter och användare och hur kan det stödja ekonomistyrningsarbetet?

1.3 Syfte

Vårt syfte är att identifiera och beskriva innebörden av begreppet Big Data samt hur de teknologier och verktyg som finns kan stödja ekonomistyrning. I syftet ingår även att sätta området i ett större sammanhang därför har vi studerat begreppet utifrån tre aktörers perspektiv.

2. Metod

2.1 Val av metod

Då vi har avsett att undersöka och förstå en ny företeelse där det tidigare inte gjorts några större mängder forskning och där resultatet är mycket svårt att förutse anser vi att det är mycket viktigt att kunna vara flexibla under arbetets gång. Vi vet inte på förhand vad som kommer vara de viktigaste delarna i vår studie utan kommer troligtvis behöva anpassa upplägg, teori och analys utefter vad vi upptäcker. Enligt Jacobsen (2002) är en kvalitativ forskningsansats att föredra för vår studie då han menar att en kvalitativ forskningsansats är mest lämplig då författaren vill skapa större klarhet i ett oklart ämne och få fram en nyanserad beskrivning av det.

Kvalitativa studier präglas även av flexibilitet i och med att man genom en frågeställning ställer upp vissa riktlinjer och hållpunkter utan att man som författare är bunden av dem på ett tidigt stadium. Uppläggning och planering kan på så vis vidareutvecklas under undersökningsfasen (Holme & Solvang, 1997).

Eftersom vårt mål är att undersöka begreppet Big Data samt hur olika aktörer anser att det kan användas för att stödja ekonomistyrningsarbetet kräver detta att vi behandlar existerande teorier kring ekonomistyrning samt hur dess innovationer sprids. Det finns mycket forskning om ekonomistyrning medan det finns mycket begränsat med forskning i ämnet tydligt relaterat till Big Data. Vi har gjort sökningar på "Big Data" i flera större databaser (Google scholar, EBSCO, Business source premier, Emerald, Mediarkivet). Då Big Data är ett såpass nytt område har träffarna med ursprung i den akademiska världen varit få. Mot bakgrund av detta har vi för att skapa en uppfattning kring området vänt oss till rapporter från de analytiker, tyckare och aktörer från utbudssidan som beskriver och behandlar området utifrån sitt perspektiv. Från våra källor kommer sedan antaganden formuleras och testas kring hur Big Data ställt mot ekonomistyrningsteori för att fastställa hur det faktiskt kan användas för att stödja ekonomistyrning utifrån den syn leverantörer, konsulter och användare har på fenomenet och dess applikationer.

Enligt Patel & Davidsson (2003) innebär ett deduktivt arbetssätt att man utifrån allmänna principer och existerande teorier drar slutsatser om enskilda företeelser man

möter. Vår studie har i enlighet med detta ett deduktivt angreppssätt genom att vi ställer teori och principer om ekonomistyrning och spridandet av dess innovationer mot företeelsen Big Data.

2.2 Ekonomistyrningsinnovationer

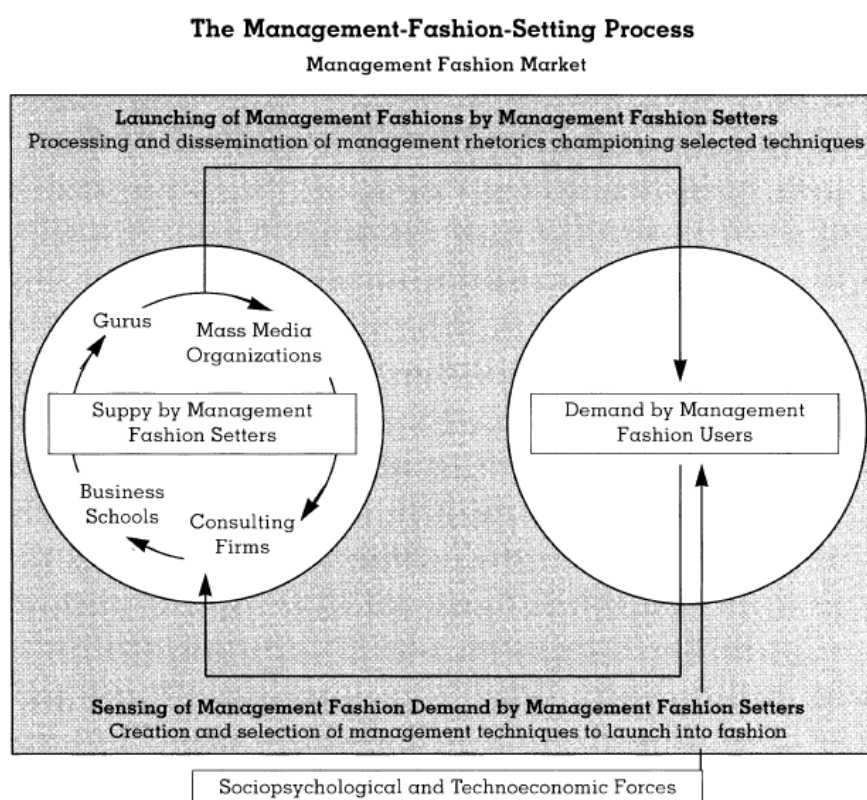
För att kunna ta avstamp i de olika aktörernas perspektiv och sätta begreppen i ett sammanhang har vi valt att se Big Data-lösningarna som ekonomistyrningsinnovationer vilket ger oss en begreppsapparat som bidrar till att beskriva de olika fenomenen på marknaden. Enligt Ax, Bjørnenak (2011) består ekonomistyrningsinnovationer av två komponenter, designelement och retorik. Designelementet är den tekniska specifikationen av innovationen medan retoriken är en diskurs som beskriver de påstådda fördelarna med att använda innovationen, och på så vis skapar den även förhoppningar. Retoriken används av utbudssidan för att övertyga möjliga användare om innovationens värde.

Det uppkommer från utbudssidan inom managementområdet ständigt nya koncept vilka syftar till att fylla de behov som finns på efterfrågesidan och lanseras till användarna. De färdiga koncepten lanseras genom paketering och spridning med till exempel media, universitet, branschorganisationer i ryggen. Med paketering menas framtagandet av den retorik vilken skall övertyga potentiella användare om konceptets nytta, nytänkande och rationalitet (Abrahamson, 1996).

Spridningen av dessa innovationer kan bero på en rad olika orsaker men efterfrågan formas i vad man kallar tekno-ekonomiska och socio-psykologiska krafter. De tekno-ekonomiska krafterna grundar sig i makroförändringar av omgivningen. Genom dessa finner företaget att de har ett prestationsgap mellan möjlig och faktisk prestation. En möjlig lösning till gap problematiken erbjuds då genom det lanserade konceptet (Abrahamsson, 1996). De socio-psykologiska krafterna handlar om hur beslutsfattare dras mot att förespråka moderna metoder i syfte att verka handlingskraftiga eller ligger i linje med de idéer som framgångsrika aktörer förespråkar (Abrahamsson, 1996).

Management fashion perspektivet fokuserar på hur innovationer växer fram genom

dynamisk samverkan mellan managers som använder sig av nya idéer och de trendsättare som paketerar dessa koncept (Birkinshaw, Hamel, & Mol). Abrahamsson beskriver med sin Management-fashion-setting process i figuren nedan, den process i vilken trendsättare ständigt producerar koncept vilka av utbudssidan presenteras som rationella och progressiva. Trendsättarna konceptualiserar och förfinar ständigt en bild de och deras följare har av vad som är den rationella vägen att gå (Abrahamson, 1996).



Figur 9. The-management-fashion-setting process

Källa: (Abrahamson, 1996)

Utbudssidan fungerar som en konceptfabrik där trendsättare bearbetar och sprider innovationer som skall fylla de behov som finns på efterfrågesidan efter de ständiga popularitetssvingningar som förekommer.

2.3 Datainsamling

För att få underlag till vår uppsats har vi använt oss av primär- och sekundärdata enligt följande:

2.3.1 Primärdata

Insamling av primärdata har skett genom två personliga intervjuer och tre telefonintervjuer. Anledningen till att vissa intervjuer gjorts via telefon har varit att respondenterna arbetat i Stockholm och vi inte haft möjlighet att resa dit för att genomföra intervjuer. Under rubriken intervjuprocessen förklaras hur vi gått till väga för att samla in kvalitativ data.

2.3.2 Sekundärdata

Vår sekundärdata består av vetenskapliga artiklar, artiklar från analytikerfirmor, böcker inom ekonomistyrning, men även så kallade “white papers” från olika leverantörer vilket är rapporter från företagssponsrad forskning. Dessutom har vi studerat kundcase från leverantörer av Big Data-lösningar. När det gäller analytikerfirmor har vi i första hand använt material från Gartner då vi haft fri tillgång till deras material genom skolan. Artiklar har i huvudsak hittats genom sökning i Google scholar samt Ekonomiska bibliotekets databaser med sökord som till exempel: Big Data, business intelligence, unstructured data, management control, management accounting och management innovations. Dessutom har referensmaterialet kompletterats med rapporter från branschaktörer inom IT.

Vår referensram har två delar då vi undersöker och jämför Big Data med ekonomistyrningsteori. Delen om Big Data är en beskrivning av området och består till stor del av information från aktörer inom branschen medan ekonomistyrningsdelen i huvudsak grundas på akademisk teori.

2.3.3 Intervjuprocessen

Vår insamling av empiri har i första hand bestått av kvalitativa intervjuer med stöd av en intervjumall då det ger möjlighet för respondenten att med vägledning av frågor fritt resonera kring vad de anser viktigt i ämnet. Enligt Holme & Solvang (1997) bör man i sammanhanget inte använda sig av standardiserade frågeformulär i kvalitativa

intervjuer på grund av att man vill komma åt intervjupersonernas egna uppfattningar och således som forskare inte bör styra för mycket. Det är istället de intervjuade som ska få styra intervjun utifrån vad de anser vara viktigt. För att genomföra denna typ av intervjuer på ett ordnat sätt rekommenderar Holme & Solvang (1997) att man som forskare använder en intervjumall. Dalen (2008) stödjer idén att semistrukturerade intervjuer med en hjälp av en intervjumall ger möjlighet att anpassa frågor och arbetsgång efter vad respondenten väljer att svara utan att tappa fokus på studiens överordnade frågeställningar och viktiga frågeområden.

Namn	Företag	Verksamhet	Roll
Simon Lidberg	Microsoft	Mjukvaruleverantör	Teknisk expert BI
Christopher Broxe	SAS Institute	Mjukvaruleverantör	Ansvarig för textanalys
Fredrik Kjellberg	Acando	Konsultföretag	Systemarkitekt
Johan Jerresand	Affecto	Konsultföretag	Practice manager
Mats-Olov Eriksson	King	Spelutveckling	Dataarkitekt

Tabell 1. Studiens respondenter

Källa: Egen

2.3.4 Urval

För att skapa en mer omfattande bild av det valda området har vi valt att inkludera tre olika intressenters perspektiv, leverantörer, konsulter och användare. Genom detta vill vi fånga upp skillnader i uppfattning kring definitioner, behov och förväntningar kring användande.

För att underlätta arbetet med att etablera kontakt med lämpliga intervjupersoner har vi använt oss av kontakter på Centrum för affärssystem vid Handelshögskolan i Göteborg och har även sökt upp personer och företag som förekommit i branschtidskrifter och seminarier rörande det aktuella ämnet. Då Big Data som begrepp är såpass nytt har vi funnit att det är en begränsad mängd möjliga respondenter som är medvetna om att de arbetar med problematiken. Vår ambition har varit att hitta personer med kompetens inom tekniken som dessutom kunnat beskriva praktisk tillämpning för att kunna sätta användandet i ett ekonomistyrningsperspektiv. Att ta hjälp via skolans kontakter var av rent praktiska skäl och vi tror inte att detta har påverkat utfallet då respondenterna verkar för olika parter på marknaden och är

bland de ledande inom sina verksamhetsområden.

2.4 Giltighetsanspråk

Intervjuer ger utlopp för personliga åsikter vilket innebär att urvalet av intervjupersoner starkt påverkar vilka svar man får, det är enligt Dalen (2008) viktigt att vara medveten om detta när slutsatser från materialet dras. Det är möjligt att vi har förbisett viss relevant information om Big Data och dess möjligheter då vi till största del utgått från de källor som vi haft fri tillgång till genom skolan. Vi har sett rapporter från andra aktörer som behandlat relevanta områden men har tyvärr inte haft åtkomst till dem.

Enligt Lincoln & Guba (1989) bildar fyra komponenter innehållet i begreppet pålitlighet för en studie. Vid kvalitativa studier är informationskällorna avgörande för studiens trovärdighet. Vi har i största möjliga mån försökt använda så aktuella källor som möjligt, men det förekommer ändå vissa äldre källor. Det är dock viktigt att vara medveten om att utvecklingen inom IT-området sker mycket snabbt. Överförbarhet handlar om hur vald indikator överensstämmer med det förhållande vi vill bedöma. Värt att reflektera över i denna studie är att ämnet är nytt och begreppen är inte klart etablerade vilket har gjort att vi har valt att låta respondenterna själva definiera använda begrepp så att vi i senare skede kunnat jämföra svar mellan respondenter.

Stabilitet avser frånvaro av slumpmässiga fel vilket i vår typ av studie kan gestaltas i vikten av att hitta respondenter med relevant kunskap i ämnet. Vidare kan stabiliteten påverkas av personliga värderingar hos respondenterna. Genom valet av personliga intervjuer har vi haft möjligheten att ställa följdfrågor och dessutom vid behov i efterhand kontakta för uppföljande frågor. Detta hade inte varit möjligt om en kvantitativ ansats med enkäter valts.

En studies bekräfningsbarhet är motsvarande dess objektivitet vilket påverkas av ursprunget på det material man använder. I vårt fall är det viktigt att vara medvetna att en del av källmaterialet härstammar från aktörer som kan ha intresse av att lyfta problematiken samt visa på nyttan med sina lösningar.

3. Referensram

Första delen av referensramen inleds med en förklaring av begreppet Big Data och följs av att det sätts i relation till traditionell BI. Vidare presenteras utbudssidans roll för framväxten av teknikdrivna ekonomistyrningsinnovationer samt hur utvecklingen av Big Data-teknologin har gått till. Andra delen rör ekonomistyrning, här presenteras teorier som senare kan relateras till Big Data.

3.1 Big Data

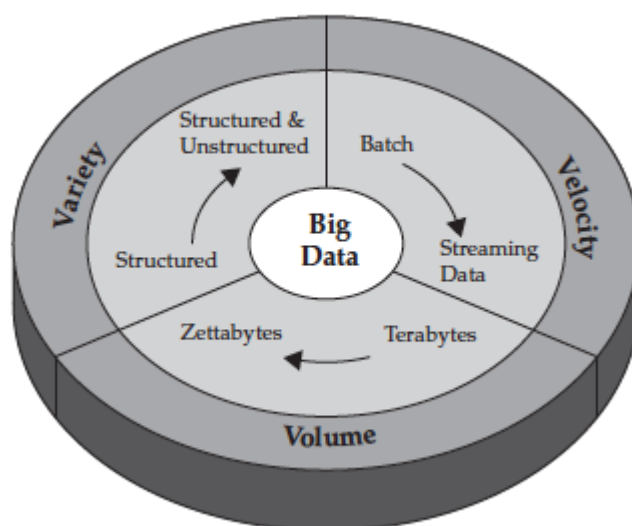
Trots att Big Data på senare tid har fått allt mer uppmärksamhet är problematiken med växande datamängder är egentligen inget nytt utan lyftes för första gången i en rapport från analytikerfirman Meta group (Laney, 2001). Rapporten noterade e-handelns framväxt vilket enligt rapporten möjliggjorde att lagra tio gånger mer data om en transaktion jämfört med traditionell försäljning. Dessutom förutspådde man att data skulle komma att genereras med ökad hastighet och att den största utmaningen för företag inom några år skulle komma att vara att hantera mängder av olika typer av data som skulle behöva omarbetas för att kunna användas tillsammans. En undersökning från Loglogic visar att många potentiella användare i dagsläget inte vet vad Big Data är och nästan skräms av vilka förändringar den nya teknologin kan innebära (Loglogic, 2012).

Den amerikanska analysfirman Gartner fastställer i en rapport från 2011 (Bitterer, 2011) att Big Data är ett av tre centrala innovationsområden inom BI. Man menar i en annan rapport (Blechar, et al. 2011) att även om begreppet fokuserar på volym ska man se termen som en början på en bredare diskussion kring fler dimensioner som blir viktiga i modern strategisk informationshantering. Gartner är bestämda i sin framtoning och menar i en rapport (Prentice, 2011) att CEO:er aggressivt bör anamma Big Data eftersom det är ett exempel på hur teknologi i framtiden kommer att skapa strategiska affärsmöjligheter och kan skapa stora vinster eller konkurrensfördelar för företagen.

Konsultfirman McKinsey väljer i en rapport från 2011 (McKinsey Global Institute, 2011) att endast fokusera på volymdimensionen men sätter inga exakta gränser för storleken utan menar att vad som klassas som Big Data varierar mellan olika branscher beroende på informationsbehov och att volymdefinitionen för Big Data

sannolikt kommer att öka med tiden.

Meta group (2001) klassade tre utmaningar som flera aktörer, bland andra TDWI (Russom, 2011) och IBM (Zikopoulos, et al. 2012) idag använder när de beskriver området. Man beskriver dessa tre områden enligt figuren nedan. De skiljer sig från dagens datahantering och tillsammans karaktäriserar de begreppet Big Data. Man talar om “de tre v:na”, volume, variety, velocity (Russom, 2011). Det förekommer även aktörer (Rhodin, 2012) som inkluderar ett fjärde V, voracity, som ska förklara den kraftigt ökade efterfrågan på data för analys från användare.



Figur 3. Big Data ur tre perspektiv

Källa: (Zikopoulos, et al. 2012)

Analysfirmorna TDWI (Russom, 2011) och IBM (Zikopoulos, et al. 2012) menar att volymen ökar snabbt och att företag idag har enstaka databaser som inte längre kan mätas i terabyte (tusentals gigabyte) utan petabyte (miljontals gigabyte) eller ännu större ses inte som ovanligt. Samtidigt ökar antalet olika datatyper som man vill ha möjlighet att lagra och analysera. Man har länge använt transaktioner från ERP, CRM, huvudbok och andra i huvudsak interna system men i och med framförallt internets massiva framväxt har mängden ostrukturerad data som finns tillgänglig ökat markant, till exempel tweets, blogginlägg och sociala medier. Variationen av olika typer av data har ökat och nu talar flera aktörer, till exempel SAS Institute, SAP och

Oracle (Dijcks, 2011) om vikten av att använda textanalys, och i viss mån även ljud- och bildanalys. Data skapas idag enligt The data warehousing institute (Russom, 2011) mycket snabbare och kommer dessutom från fler källor än tidigare.

Utvecklingen har gått från gammal batchhantering där stora mängder transaktioner lagrades inaktiva tillfälligt för att senare samlas och köras klumpvis vid ett eller ett fåtal gånger per dygn till att idag i många fall uppdateras i realtid.

3.1.1 Relation till Business Intelligence

Baars, Kemper (2008) har arbetat fram ett ramverk för strukturen hos traditionella BI-lösningar som vanligtvis används för analys, rapportering och visualisering av i huvudsak strukturerad data från interna källor. De menar att en BI-lösning består av tre olika hierarkiska lager: datalager, logiklager och tillgänglighetslager. Datalagrets uppgift är att lagra strukturerad data, detta görs vanligtvis i data warehouse på grund av att datan är organiserad i rader och kolumner. Detta lager ska också kunna hantera ostrukturerad information. Datalagret laddas från både interna (exempelvis ERP, CRM) och externa (exempelvis köpta marknadsanalyser) källor efter att ha omformats i en tvättningsprocess där olika strukturer översätts till en enhetlig, denna process kallas ETL (Extract, Transform, Load). Logiklagret tar sedan hand om analys- och rapporteringsfunktionalitet som kan bestå av exempelvis att skapa och sprida rapporter eller analyser både i återkommande eller adhocform. Här kan även prestationsmätning, till exempel balanserat styrkort, göras på den data som rapporteras. Tillgänglighetslagret är ett presentationsgränssnitt som ska underlätta för användarna att komma åt funktionerna i logiklagret enligt fördefinierade ansvarsroller och befogenheter. Vanligtvis levereras det sista lagret som en portaltjänst med tonvikt på visualisering med till exempel diagram och liknande.

TDWI (Russom, 2011) hävdar att traditionella BI-lösningar inte räcker för att hantera den ökade komplexiteten och volymen hos riktigt stora och komplexa datamängder. Man menar också att risk finns för att en ETL-process rensar bort värden som förefaller vara felaktiga eller av för låg kvalitet, vilket kan göra att man kastar guldgruvorna som potentiellt kan ge just de viktiga insikter som Big Data kan tydliggöra. TDWI har gjort en undersökning där 42 % av användarna klagade på att existerande analysplattformar inte kunde hantera tillräckligt stora datavolymer, 29 %

upplevde att data laddades för långsamt och 24 % tyckte att svarstiden var för lång (Russom, 2011).

3.1.2 Strategisk och operativ BI

Skirletz (2002) menar att BI kan användas på två sätt, operativt eller strategiskt, enligt bilden nedan. Att använda det operativt innebär att standardisera, organisera och leverera information och avvikelser till rätt person för att hjälpa dem att göra sitt arbete. Det praktiska fokuset på att lösa operativa problem och underlätta arbetet begränsar däremot värdet på BI-användandet enligt Skirletz genom att endast leverera det de anställda i förväg vet att de behöver.



Figur. 4 Business Intelligence ur operativt och strategiskt perspektiv.

Källa: (Skirletz, 2002)

Skirletz menar att företag har mycket att vinna på att utveckla sitt BI-användande från operativt till strategiskt. Nyckeln till det är enligt honom att ändra synsättet på vilken information man kan få fram. Om man istället söker efter nya typer av svar rör man sig mot ett strategiskt användande där man kan öka vinst, hålla kostnader nere, förhindra förluster och söka nya möjligheter. Användandet av BI begränsas enligt

författaren endast av förmågan att samla användbar data, fråga intressanta frågor baserat på datan för att få insikt och agera effektivt på resultaten.

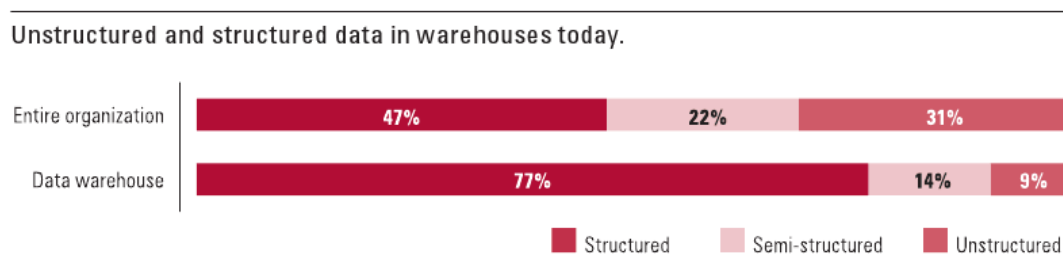
3.1.3 Var skapas informationen

I och med informationsteknologins spridning skapas data på allt fler håll i infrastrukturen. Idag talas det om att anpassa tekniken för att möjliggöra analys av fler typer av data. Oracle (Dijcks, 2011) benämner nuvarande typer som kan användas för analys i tre kategorier. Först och främst finns traditionell affärsdata vilket inkluderar kundinformation från CRM-system, transaktionsdata från ERP, webshop eller huvudbok. I takt med att fler och fler maskiner och processer använder IT i större utsträckning har också maskingenererad eller sensordata kommit in i bilden. Detta kan till exempel vara samtalshistorik i callcenters, GPS-data, "smarta mätare" inom produktion samt loggar från olika typer av utrustning och handelssystem. Slutligen nämner Oracle också social data vilken oftast är textbaserad och sprids på internet till exempel genom sociala medier som Twitter och Facebook.

Gartner (2011) (Sallam, et al. 2011) beskriver en förändrad efterfrågan av BI-lösningar hos kunder, många vill nu använda sina BI-lösningar på nya områden. Förekomsten av analys av sociala medier har ökat efterfrågan på analys av stora ostrukturerade datamängder, i huvudsak textbaserade, att kunna tolka meningen av dessa.

De beskriver även att det finns stora mängder ostrukturerad information inom organisationers väggar, som även det kan användas för analys. Däremot råder det delade meningar om hur fördelningen är mellan strukturerad och ostrukturerad information. Merill Lynch uppskattade redan 1998 att 80 % av den användbara affärsinformationen förekom i ostrukturerad form. Den siffran har sedan dess spridit sig och blivit vedertagen inom branschen. En undersökning gjord av TDWI (Russom, 2007) med 370 respondenter kom fram till andra resultat, man såg snarare att någonstans kring 50 % av organisationers information var ostrukturerad. Någon exakt siffra är svår att komma fram till då olika definitioner på typer av data kan göra det svårt att jämföra korrekt. TDWI (Russom, 2007) undersökning redovisad på bilden

nedan var däremot 77 % av dåtidens datawarehouse laddade med strukturerad data medans endast 23 % var ostrukturerad, här verkar Merill Lynchs uppskattning däremot stämma ganska bra.

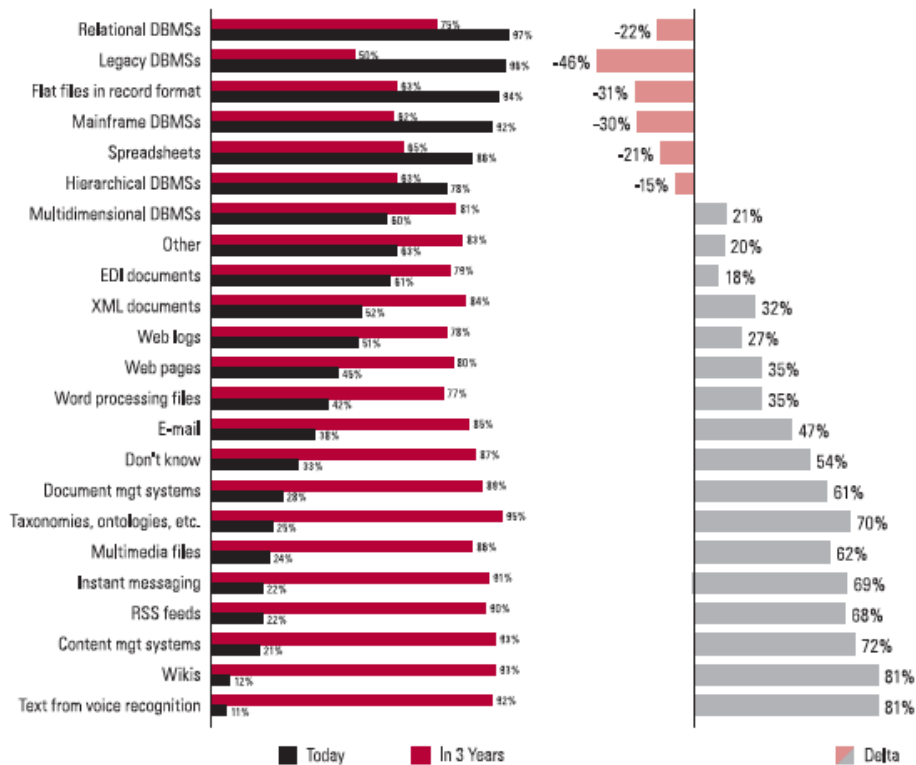


Figur 5. Andel ostrukturerad och strukturerad data i organisationer jämfört med datawarehouses 2007 enligt TDWI.

Källa: (Russom, 2007)

I samma undersökning frågade man respondenterna om vilka specifika datatyper man hade i sina datawarehouse och bad även respondenterna uppskatta hur det skulle se ut tre år senare, 2010. Man kan urskilja en tydlig trend i att det vid tillfället för undersökningen var stort fokus på traditionella databastyper och att detta i framtiden spåddes avta till förmån för till exempel multimediafiler, e-post och XML/RSS (semistrukturerad information).

Which types of data and source systems feed your data warehouse? (Select all that apply for both today and in three years.)



Figur 6. Beskriver data i datawarehouse ursprung efter källor.

Källa: (Russom, 2007)

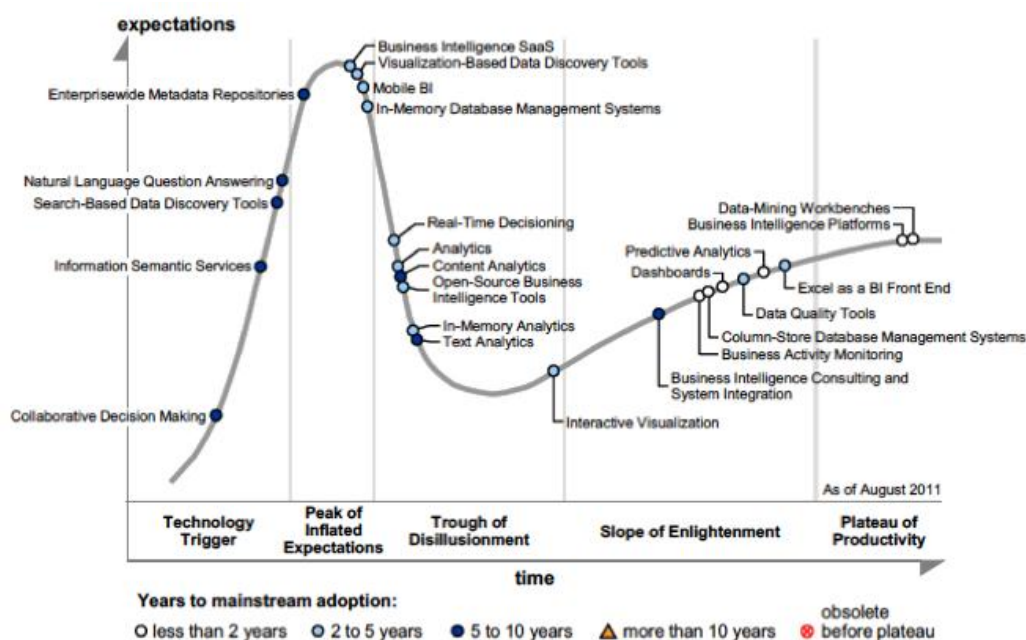
3.1.4 Design och paketering

På utbudssidan av innovationer finns ett antal tongivande parter som sprider forskning, rådgivning och anordnar konferenser inom informationsteknologi och systemstöd för näringslivet. Enligt Pollock, Williams (2010) speglar dessa aktörer inte bara rådande marknadssituation inom ett visst område utan bidrar aktivt till att forma denna. Författarna menar exempelvis att det amerikanska analysföretaget Gartner formar innovationer genom att man i förväg namnger och definierar framväxande teknologier. Istället för att i efterhand anpassa sin definition efter vad som presenteras av leverantörer förväntar man sig att dessa anpassar sig till redan framtagna definitioner. Analysföretag inom IT-sektorn tas upp som extra tydliga exempel på en maktfaktor men det finns enligt författarna liknande situationer även i andra branscher, framförallt där det råder stor osäkerhet och utvecklingstakten är hög.

En av de största inom analys av IT-marknaden är det redan nämnda amerikanska företaget Gartner som bedriver forskning och rådgivning riktat till chefer och senior personal inom i huvudsak större organisationer och statliga institutioner (Gartner

Inc.). Ytterligare aktörer inom analys och rådgivning är Forrester research och TDWI om även de bedriver forskning och rådgivning för organisationer som är intresserade av teknologitveckling (Forrester Research Inc.) (The Data Warehouseing Institute). Utöver dessa finns även globala konsultfirmor, exempelvis McKinsey, och systemleverantörer (exempelvis Microsoft, SAP, Informatica, Oracle) som även de är inflytelserika på området och sponsrar forskning samt presenterar rapporter över utvecklingen inom IT-området.

För att presentera hur olika teknologier rör sig mot bred adoption på marknaden använder Gartner en egenutvecklad modell kallad "Hype cycle". Denna modell har man använt sedan 1995 (Fenn & Raskino, 2011) och syftar till att spegla hur man anser att nya teknologier går igenom fem faser innan förväntningarna på ny teknologi blir rimliga och lösningarna når den breda massan och användningen är etablerad (Fenn & Raskino, 2011). Nedan följer exempel på områden med koppling till Big Data som Gartner presenterade i sin hype cycle för BI 2011.



Figur 7. Gartners Hype Cycle för Business Intelligence 2011

Källa: (Bitterer, 2011)

Content analytics bearbetar innehåll av flera olika slag och använder beteendemönster när det gäller konsumtion av innehåll för att härleda svar till specifika frågor.

Exempel på typer av innehåll kan vara alla former av text som dokument, bloggar, nyheter, konversationer med kunder och diskussioner i sociala medier. Möjliga

analysmetoder av detta material kan vara textanalys, talanalys samt beteendeanalys och kan användas inom flera olika funktioner. Det skulle till exempel kunna bidra till att identifiera viktiga kunder eller upptäcka problem med produkter eller tjänster. Man kan även analysera konkurrenters aktiviteter och kunders reaktioner på nya produkter. En stor fördel med tekniken är att man kan automatisera komplicerat och tidskrävande analysarbete på ett lätthanterligt sätt som annars hade gjorts manuellt. Gartner bedömer att området bör ha bred spridning om fem till tio år (Bitterer, 2011).

Textanalys innebär att härleda information från olika källor av text till exempel tweets eller anställdas noteringar om kundsamtal. Det kan antingen summera, förklara eller utreda möjliga orsaker till olika händelser och intresset har enligt Gartner ökat bland användare vilket har lett till att fler leverantörer har utvecklat tjänster för ändamålet. Den största effekten av textanalys tros komma när man lyckas kombinera det med traditionell strukturerad data och kan då ge en helhetsbild som vanliga BI-lösningar inte kan ge. Förhoppningen från användare bedöms fortfarande vara för hög i relation till dess faktiska värde, men spridningen tros bli mycket stor när förväntningarna blir realistiska. Gartner bedömer vara moget om fem till tio år.

Tillsammans med prognoser ses sentimentanalys som ett av de hetaste analysområdena inom BI och analys (Herschel, 2010). Sentimentanalys är ett område inom textanalys vilket syftar till att avgöra huruvida en text är positivt eller negativt inställd till ett sökord. Den här typen av analyser används främst till att följa upp hur produkter och varumärken tas emot i de sociala medierna i syfte att få marknadsunderlag (Pant, Lee, & Vaithyanathan, 2002).

3.1.5 Teknik

För att kunna hantera och genomföra analys av stora, semi- eller ostrukturerade datamängder har det i vissa fall krävts att man söker nya lösningar vid sidan av den traditionella relationsdatabasen. Därför har fokus till stor del vänts bort från de konventionella SQL baserade databaslösningarna (King, 2011). Tidig Big Data kretsar ofta kring NotOnlySQL-miljöer eller NoSQL, vilket är ett samlingsnamn för en rad olika tekniker som inte bygger på den klassiska relationsdatamodellen eller använder sig av SQL queries (Neubauer, 2010).

Många av de Big Data lösningar vi idag ser ute hos användare och de som växer fram i de stora leverantörernas produktpaket bygger på Apache software foundations Hadoop ramverk. Hadoop nämns allt mer i sammanhang med Big Data och ses av analytikerfirman Gartner som en av de tio teknologitrender som med mest sannolikhet kommer att få stor betydelse under 2012 (Casonato, 2011)

Hadoop är ett projekt sprunget ur Apachecommunityn och ramverket är open source-mjukvara för att dela upp stora krävande beräkningar på flera vanliga enheter och innehåller främst filsystemet HDFS och ett MapReduce-ramverk (Apache Software Foundation, 2012) Inspirationen till Hadoop är hämtad ur Googles MapReduce och Google filesystem. MapReduce är en teknik vilken från början skapades av Google (Lämmel, 2007) för att kunna hantera stora komplexa datamängder i parallella processer effektivt.

Då Hadoop ramverket är open source innebär det att vem som helst kan ladda hem programvaran gratis och använda eller hjälpa till att vidareutveckla ramverket (Apache Software Foundation, *What is Hadoop?*, 2012). Hadoop används idag på en rad större företag främst inom IT branschen, på användarlistan finns bland annat Amazon.com, Apple, Facebook, Fox, IBM, Twitter (Apache Software Foundation, 2012) Men spridningen av teknologin utanför sektorn har varit begränsad, delvis på grund av att lösningarna inte varit särskilt användarvänliga och därför i huvudsak kretsat kring företag med mycket stor kompetens inom programmering.

Det är fram tills nu i huvudsak mindre aktörer i branschen som erbjudit paketerade lösningar baserat på Hadoop men de stora aktörerna ligger i startgroparna med sin paketering (Microsoft , 2011). I flera fall rör det sig i dessa paketeringar om att integrera Hadoop i de redan befintliga produkterna. Genom så kallade connectors kan användarna flytta data mellan Hadoop och SQL och på så vis integrera den nya informationen och vice versa.

Analysfirman Forrester släppte i februari 2012 en rapport där de ger sin syn på marknaden för Hadoop-lösningar (Kobielus, 2012). Enligt dem ska man som användare inom området vara försiktig då det fortfarande är en ny teknologi och

marknaden fortfarande är omogen. Man hävdar att flera paketerade lösningar ännu är omogna och det finns en stor variation i deras funktionalitet, prestanda och flexibilitet. Forrester spår även att lagringen av stora datamängder inom Big Data kommer att kompletteras av, eller helt domineras av molntjänster under de kommande åren.

Samtidigt är det viktigt att påpeka att det händer saker inom området även vid sidan av Hadoop. Flera av de stora utvecklarna arbetar med tekniker för hantering av relationsdatabaser vilken skall göra dessa kraftfullare ur ett Big Data perspektiv. Genom introduktion av tekniker såsom parallellitet (arbetet delas upp på flera maskiner) och in-memory (informationen laddas i arbetsminnet) förbättras analysmöjligheterna (Bitterer, 2011).

3.2 Ekonomistyrning

Ekonomistyrning eller Management Control Systems definieras på många sätt i litteraturen. Dessa innefattar olika koncept vilka sorterar de kontrollmekanismer som skall styra organisationens verksamhet i rätt riktning (Malmi & Brown, 2008).

Även Merchant & Van der Stede (2007) har en liknande definition där de menar att ekonomistyrningen ytterst handlar om att kontrollera mäta och styra beteende. Ekonomistyrningen har genom åren utvecklats från att enbart innefatta formella och finansiella mått till att även innefatta bitar såsom icke-finansiell information, relationer och social kontroll. Övergripande syftar dock ekonomistyrningen till att stödja företaget i arbetet med att uppnå förbestämda mål. I begreppet ingår planering, genomföring, uppföljning, utvärdering och anpassning av företagets verksamhet för att uppnå dess bestämda ekonomiska mål (Ax, Johansson, & Kullvén, 2009).

Begreppet ekonomistyrning innefattar alltså olika styr- och kontrollverktyg. Författarna nämner också att fokus historiskt legat på finansiella mål medan utvecklingen går mot att icke-finansiella mål blir allt viktigare. Nöjda kunder och medarbetare anses bidra till att uppfylla finansiella mål. Samuelsson & Olve (2001) definierar de styrmedel som finns enligt tre kategorier där en, formella styrmedel, handlar om att fastställa och följa upp den strategi och planering som utarbetats från

affärsidén. En förutsättning för att dagens ekonomistyrning skall fungera är möjligheten att tillhandahålla den information som krävs från underliggande system på ett tillförlitligt sätt. Här fungerar insamlad information och BI verktygen som indata till ekonomistyrningen. Genom att tillhandahålla rätt information kan bitar som planering, uppföljning, utvärdering få ett bättre underlag.

I dagens organisationer är det en ständig kamp att vinna och behålla konkurrensfördelar. För att göra detta krävs ofta snabb förändring även av ekonomistyrningen för att anpassa och effektivisera arbetet. Michael Porter och Victor E. Millar skrev redan 1985 om hur mängden information var kraftigt tilltagande. Ökad tillgång och minskande kostnader att samla in, hantera och överföra information leder enligt Porter & Millar till nya möjligheter och affärsmodeller. Författarna pekar på hur informationsteknologin leder till skiften mellan konkurrenter inom branscher. Att hantera informationen på ett bra vis kan leda till konkurrensfördelar, främst på kort sikt då aktörer inom branschen tenderar att följa de som är framgångsrika och det är svårt att monopolisera på IT-lösningar.

3.2.1 Management control package

Malmi och Brown (2008) talar om ekonomistyrning i termer om ”management control package” vilket illustreras i figuren nedan. De menar att man under lång tid beskrivit kontroll och styrning som en sammansättning av en rad kontrollmekanismer vilka samverkar och betonar vikten att se till paketet som helhet snarare än hur de enskilda kontrollmekanismerna fungerar var för sig. De betonar vikten att inte enbart se till de finansiella delarna när man studerar styrsystem utan även väga in delar såsom kulturell styrning och administrativ kontroll och hur detta bidrar till organisationens måluppfyllnad. Det är alltså viktigt att nya mekanismer studeras som en helhet i systemet snarare än hur den verkar isolerat. De mekanismer som ingår i paketet definieras som kontroll och styrning av kultur, planering, cybernetik-, belöning- och administration. De olika delarna har adderats till styripaketet i olika stadier då den utvecklats (Malmi & Brown, 2008).

Cultural Controls						
Clans		Values			Symbols	
Planning		Cybernetic Controls				Reward and Compensation
Long range planning	Action planning	Budgets	Financial Measurement Systems	Non Financial Measurement Systems	Hybrid Measurement Systems	
Administrative Controls						
Governance Structure		Organisation Structure			Policies and Procedures	

Figur 8. Malmi & Browns Management control system package

Källa: (Malmi & Brown, 2008)

Malmi & Brown (2008) tar även upp policys och procedurer som handlar om att specificera processer och beteende i en organisation genom att sätta upp regler och arbetssätt. Övervakning av anställdas handlingar och standardisering av processer ska säkerställa att anställda handlar effektivt och i organisationens intresse.

Den cybernetiska styrningen är dock särskilt intressant i sammanhanget. Den beskrivs av Malmi & Brown enligt fem egenskaper:

- Den innehåller åtgärder som gör det möjligt att mäta och kvantifiera underliggande fenomen, aktiviteter eller system.
 - Det finns normer och prestationsmål
 - Det finns möjlighet till jämförelse och återkoppling av prestationer
 - Variationsanalys utifrån feedback
 - Möjligheten att ändra systemets beteende eller underliggande aktiviteter
- (Malmi & Brown, 2008).

Simons (2000) relaterar resultatstyrning och mätning till implementering av nya strategier vilka skall leda till att företagets mål nås. Han exemplifierar sådan strategiimplementering till att minska kostnader, utveckla IT-användande samt öka kundnöjdhet. För att målen skall bli angripbara krävs att de går att kvantifiera eftersom man ska kunna mäta måluppfyllelsen.

Porter (1996) menar att det är ett missförstånd att inre effektivitet, dvs. "att göra saker på rätt sätt" (Ax, Johansson, & Kullvén, 2009), är tillräckligt för ett företag för att bli framgångsrikt. Arbetsätt som effektiviserar de egna processerna kan oftast kopieras av andra aktörer och leder till att den absoluta produktiviteten, men inte den relativa, ökar. Alla springer samma lopp snabbare och snabbare. Vad Porter istället menar är att strategi består av att använda unika aktiviteter som skiljer sig från konkurrenternas och att våga kompromissa för att undvika ett för brett kunderbjudande.

Porter (1996) beskriver också hur det är extra svårt att skapa en hållbar strategisk position i branscher där teknologisk utveckling är mycket snabb. Under dessa förhållanden är osäkerheten stor över kunders preferenser, vilka teknologier som kommer att bli etablerade och framgångsrika vilket gör att aktörer av rädsla för att hamna efter tenderar att imitera varandra vilket i längden inte är hållbart. Porter (1996) menar att detta ofta sker under längre perioder i högteknologiska branscher än andra.

4. Empiri

Första delen av empirin består av fyra case som syftar till att ge verkliga exempel på praktisk användning av Big Data-teknologi. Andra delen består av de intervjuer vi gjort under arbetets gång och går igenom respondenternas syn på begrepp, användning/designelement, retorik/behov och avslutas med en sammanfattning utformad som en tabell.

Case 1: Vestas Wind Systems

Det danska vindkraftföretaget Vestas är världsledande på vindkraftturbiner. För att maximera effekt samt nyttjandegrad av dessa är placeringen mycket viktig. Det är inte en enskild variabel som utgör en bra plats, utan flera faktorer spelar in. Under 2011 genomförde företaget en stor satsning på teknik från dataföretaget IBM för att möjliggöra omfattande analys av strukturerad och ostrukturerad data som väderrapporter, tidvattenfaser, geo- och sensordata, satellitbilder, kartor över avskogning och väderprognoser i syfte att kunna ta smartare beslut om sina investeringsprojekt (IBM media relations, 2011). Genom analysen kan de inte bara bestämma vilken plats och riktning som lämpar sig bäst utan även prognostisera hur mycket elektricitet potentiella vindkraftsprojekt kan generera och på så sätt kunna göra bedömningar om potentiella investeringar är motiverade. De kan dessutom kombinera sensordata med väderdata och följa varje blad i kraftverken vilket möjliggör planering av underhåll i rätt tid och genomföra det vid tillfällena produktionen ligger på en låg nivå vid till exempel turbulens. Till sin hjälp för hantering av data och analys har de Firestorm, en mycket kraftfull superdator med 14664 stycken kärnor och en lagringskapacitet på 2,6 petabyte vilken använder IBM big insights som är IBMs enterprisepaketering av Hadoop (Vestas Wind Systems, 2011).

Case 2: Station Casinos

Det finns många casinos i Las Vegas, för att kunna stå sig i konkurrensen bestämde sig Station Casinos att lära känna sina kunder bättre genom att införa realtidsupptagning och integration av data från över 500 olika källor på deras casinon där samtliga händelser loggas. Man samlar in information om allt ifrån kunders casinospel till hotellvistelser och restaurangförsäljning. Genom att sedan

analysera kundernas beteenden inne på anläggningarna kan företaget skapa sig en bild hur varje kund spenderar sin tid i lokalerna och sedan skraddarsy erbjudanden och ta strategiska beslut baserat på informationen i syfte att maximera värdeskapandet (Informatica , 2011). Företagets BI-ansvariga menar att tekniken ger dem möjlighet att anpassa servicenivån efter kundvärde. Detta förbättrar verksamhetens rörlighet och ger konkurrensfördelar då de kan förutse kundernas behov. Genom analys av informationen kan de avgöra vilka kunder som är de mest lönsamma samt vilka åtgärder som leder till ett ökat kundvärde. Arbetet har lett till att man kunnat öka antalet kundsegment från 40 till 160 stycken att anpassa sina erbjudanden mot. För att genomföra detta har de en rad verktyg från leverantören Informatica. De hanterar en rad olika datakällor och innehåller både strukturerad och ostrukturerad information vilken adderas och hanteras i loggarna i realtid. (Informatica , 2011).

Case 3: AFA Försäkring

AFA försäkring är ett svenskt försäkringsbolag som via kollektivavtal försäkrar anställda inom privat, kommunal, landsting och regional sektor. (AFA Försäkring, 2011). Alla som anmäler en skada, får i textfält beskriva orsaken till skadan, händelseförloppet och vilka konsekvenser den har medfört. Försäkringsbolaget har som ansvar att analysera orsaker till skador, sjukdomsfall och dess konsekvenser för att rapportera till arbetsmarknadens parter som kan genomföra skadeförebyggande åtgärder. Detta innebär stora mängder insamlad ostrukturerad data i rapportform som skall gås igenom och analyseras efter samband i 10000-tals dokument. För att effektivisera arbetet använder man sig av ett textmining-verktyg från SAS Institute (SAS Institute, 2009). Rent praktiskt går det till så att när en anställd blir skadad på sin arbetsplats skall det rapporteras vad som föregick, hur det gick till och vad konsekvensen av olyckan blev. De här filerna har man tidigare försökt göra rapporter på där man försökt förstå vilka olyckor det är som orsakar de värsta följderna samt inom vilka områden man kan göra förbättringar. Tidigare hade man anställda som läste incidentrapporterna manuellt och kodade dessa för att möjliggöra analys. Med textanalysen har de fått fram mycket snabbare och mer exakt resultat och även kunnat lära sig nya saker som man inte haft ett kodningsschema för. Försäkringsbolaget kan med hjälp av textanalysen se mönster och upptäcka överrepresentation av vissa typer

av skador (Broxe, 2012). På sikt vill man även kunna använda informationen förebyggande. De vill kunna prognostisera risker och allvarlighetsgrader för olika yrkesgrupper baserat på statistisk analys av materialet. Bakgrunden till att AFA försäkring började arbeta med textmining är enligt Michel Normark, chef för analys och försäkringsvillkor framför allt att de ville höja kvaliteten på analyser och beslutsunderlag. Samtidigt var det ett sätt att sänka administrationskostnaderna genom bortrationaliserade arbetsmoment (SAS Institute, 2008).

Case 4: King

King är världsledare inom små sociala spel på webben, de har idag över en miljard speltillfällen varje månad inom sina affärsområden web-, social- och mobile-gaming. King är helt datadrivna och har i princip ingen annan information än den som finns loggad från spelen. Tillgången till data är mycket viktig och nödvändig för allt från finans till produktutveckling (Eriksson, 2012). King använder sig av en egen lösning där de sparar samtliga händelser som sker kring spelandet i separata textfiler vilka sedan kan samköras i analys. Detta skapar stora datamängder som måste hanteras. Problematik uppkom när företaget växte och skulle slå sig in på social gaming. Deras gamla databaser klarade inte av antalet händelser och datamängderna blev allt större. King kom då att utveckla en egen lösning med teknik baserad på Hadoop och verktyget Hive för att ställa analysfrågor. Förutom att de nu kan hantera alla transaktioner fick de även möjlighet att kunna mäta och analysera minsta lilla händelse. Detta har öppnat dörrar gällande allt ifrån produktutveckling till lönsamhetsberäkningar. Genom att mäta prestanda, second day retention (andel återvändande spelare), hur långt spelarna avancerar och vilka funktioner de använder kan produktcheferna få viktigt input i hur de skall utveckla sina spel (Eriksson, 2012). Märker en produktchef att hans eller hennes spel har dålig second day retention kan han eller hon gå in och direkt få ledtrådar om orsaken vilken till exempel skulle kunna vara att prestandan är för dålig vilket gör att spelet inte flyter för spelarna och kan då åtgärda detta.

Analys av betaltransaktioner ger redovisningsunderlag och information för finansavdelningen. I och med att all information finns sparad och enkelt kan tas fram vid behov möjliggör den enligt datarkitekten Mats-Olov Olsson för stora möjligheter till kreativt användande.

4.1 Intervjuer

4.1.1 Introduktion

Leverantörer

Microsoft

Microsoft är ett globalt datorföretag och idag världens största aktör inom programutveckling. Microsoft grundades 1975 av Paul Allen och Bill Gates och har kommit att vara en av de mest inflytelserika aktörerna i informationsteknologins spridning. Företaget har idag strax under 100 000 anställda och deras produkter är mycket spridda (Microsoft News Center, 2012). Simon Lidberg, teknisk specialist på beslutstöd och databaser beskriver hur Microsoft idag arbetar på en bred front när det gäller Big Data. De har forskning och utveckling kring alla större områden och erbjuder idag paketerade verktyg för såväl olika typer av analys som hantering av stora datamängder.

SAS Institute

SAS Institute är idag världens största privatägda mjukvaruföretag med över tolv tusen anställda på kontor över hela världen (SAS Institute, 2012). Företaget grundades i mitten på 70-talet i USA av en universitetsprofessor i statistik och fokuserade inledningsvis enbart på statistiklösningar men har under åren växt inom andra närliggande områden. Idag beskriver Christopher Broxer, som är globalt ansvarig för deras verktyg för textanalys, att man erbjuder lösningar inom dataintegration, databaser, statistisk analys, ostrukturerad analys och rapportering/visualisering.

Konsulter

Acando

Acando är ett konsultföretag som arbetar med verksamhetsförbättring genom informationsteknologi. Man har tjänster inom management, strategisk IT, BI, affärssystem och arbetar med allt från systemutveckling till rådgivning. Acando har idag cirka 1000 medarbetare i fem europeiska länder. Fredrik Kjellberg arbetar som systemarkitekt inom Java och är även med i en arbetsgrupp för BI. Han har idag även

som uppgift att titta på området Big Data och hur det skulle kunna användas ute hos kunder.

Affecto

Affecto är konsulter inom informationshantering med inriktning på rådgivning och implementation av beslutsstöd, performance management och dataintegration. Totalt är man runt 900 anställda varav cirka 100 finns i Sverige. Johan Jerresand är Practice manager vilket innebär arbete med verksamhetsutveckling, bland annat kring hur nya lösningar skall levereras till kund.

Användare

King

King (tidigare Midasplayer) är ett spelföretag som startade 2003 genom webbplatsen midasplayer.com. De driver sidan där spelare satsar pengar och spelar mot varandra i skicklighetsbaserade spel. På detta tar King sedan en rake, provision, på pengar som satsas. Utöver detta är man även verksam inom social gaming och utvecklar spel för Facebook och liknande. Inom detta område är affärsmodellen annorlunda och handlar om att sälja premiumtjänster såsom ingame-items och extraliv. King håller dessutom på att börja med mobile gaming. Företaget har expanderat kraftigt och har idag ett hundratal anställda på kontor i London, Polen, Rumänien, Finland, San Fransisco, Stockholm och Malmö. Mats-Olov Eriksson är idag data arkitekt på King och har tidigare arbetat i en roll som BI developer.

4.1.2 Begreppet Big Data

Leverantörer

Simon Lidberg på Microsoft beskriver hur själva termen Big Data har sitt ursprung i företag som driver stora webbfarmar såsom Yahoo och Google. Dessa hanterar stora mängder data och har ett intresse av att veta vad som sker på deras servrar. Begreppet i sig är svårdefinierat och för att beskriva det väljer Lidberg att titta på Gartner och Forrester som pratar om kriterier i form av ”tre eller fyra V”, volume, velocity, variety. Han menar att hanteringen av data i storlekar över petabyte är svårt, ofta ligger problematiken i samkörningen. Det finns även en problematik när det gäller att hantera hundratusentals händelser vilka i vissa fall kan ske under väldigt kort tid. En

central fråga är hur man skall hantera analys av data i olika format. De kundrepresentanter han möter i sitt arbete är främst marknadsförare som vill veta vad som sker kring företaget på internet. Men även olika former av analytiker, experter på algoritmer och statistik som utför analyser vilka i förlängningen kan leda till stöd i strategiska beslut.

Christopher Broxe på SAS Institute talar i Big Data i termer av hur datamängden i världen idag ökar, både vad det gäller strukturerad och ostrukturerad data såsom video bilder, ljud, text, sociala medier. Frågan är hur hård- och mjukvara kan stödja analys av dessa stora datamängder och ge användarna relevant information. Användandet kan leda till att ge bättre beslutsunderlag men Broxe menar även att verktygen kan användas för att sänka kostnader då många företag idag arbetar med att analysera ostrukturerad information även om det inte själva tänker på att de gör det. Ett exempel är att anställda ofta sitter och läser igenom dokument eller liknande. Detta kan vara ganska kostsamt i jämförelse med att automatisera de här processerna vilket i förlängningen även skulle kunna leda till bättre beslut. Broxe nämner även en annan sida som gäller sociala medier och det som skrivs på nätet om ens produkter och varumärken vilket är näst intill omöjligt att hålla reda på. Här menar han att det finns stora förtjänster att ha ett system som kan förstå detta då man tidigt kan fånga upp avvikelser.

Konsulter

Fredrik Kjellberg på Acando talar om Big Data som i regel ostrukturerade datamängder från ca 1 terabyte och uppåt. Storleken gör att det tar för lång tid att analysera på konventionellt vis. Han menar att syftet med all datainsamling inte alltid är helt klart vid insamlingsögonblicket men att det senare kan användas till att analysera trender och samband. Dessutom är lagring billigt och blir allt billigare. Det krävs ingen speciellt avancerad hårdvara utan bara ett antal vanliga hårddiskar som enkelt kan bytas ut. Teknologin ställer heller inga höga krav på processorkraft utan man löser det genom att fördela arbetet på ett antal stationer. Kjellberg menar då att de låga kostnaderna kan göra det gynnsamt att spara information vilken man kan komma att ha nytta av vid analys i framtiden allt eftersom datalagret växer fram kan man hitta nya användningsområden för den information organisationen har.

Johan Jerresand bedömer volym som centralt när man ska definiera Big Data men nämner även velocity och variety. På Affecto lägger man även till ett fjärde perspektiv, komplexitet, som är ett resultat av de tre första perspektiven. Jerresand menar samtidigt att problematiken egentligen inte är ny, utan att till exempel telekom och handel länge samlat på sig stora mängder data men inte använt den i någon större utsträckning. Syftet med att använda Big Data är att det finns stora affärsmöjligheter. Jerresand menar att många industrier står inför en revolution där man kan till exempel anpassa prissättning löpande. Han exemplifierar detta genom att nämna de smarta elmätare som nu introduceras vilka kommer att rapportera elförbrukning mer frekvent än tidigare. Det finns med hjälp av Big Data möjligheter att kunna differentiera sig och anpassa sitt kunderbjudande på ett nytt sätt.

“De som inte anammar detta, de kommer inte ha ett existensberättigande i framtiden. I det korta perspektivet handlar det om konkurrensfördelar och det långa loppet om existens. Man behöver inte nödvändigtvis börja använda det imorgon, men man kan inte vänta väldigt många år”.

Användare

Mats-Olov Eriksson på King menar att det finns många spridda tankar om vad Big Data innebär. Just hos King menar han att det är antalet transaktioner som är avgörande men generellt så är det helt enkelt när det blir för mycket data för en traditionell SQL-lösning, när det inte går att lägga till ytterligare en shard utan måste hitta en annan lösning. Då går man vidare till en NoSQL-lösning och vill på något sätt distribuera arbetet mellan flera datorer. Det är enligt Eriksson svårt att sätta exakta siffror på när skiftet bör ske. Vidare berättar han att det även finns de som vill lägga lite mer komplexitet till begreppet till exempel när det finns flera olika källor och man använder datan på ett smart sätt ur ett makroperspektiv.

4.1.3 Användning och designelement

Leverantörer

Microsoft arbetar enligt Lidberg idag med Big Data utifrån två vinklar. Dels den stora strukturerade databasplattformen utifrån SQL-server där man har produkten Parallell data warehouse vilket är en så kallad “massive parallell process”-lösning där man använder flera servrar parallellt för att hantera data och frågor. Den andra vinkeln är

analys av ostrukturerad information där man arbetar som en del i Hadoop-projektet. Microsoft håller idag på att paketera en egen distribution av Hadoop och en del kringtjänster såsom möjligheten att utföra MapReduce-jobb i JavaScript vilket man hoppas skall göra det mer tillgängligt för användare då det är betydligt fler tekniker som behärskar den typen av script. Tidigare har Hadoop-lösningar krävt att utvecklarna kan skriva egna javaprogram vilket är något som långt ifrån alla behärskar. Microsoft erbjuder även verktyg för andra områden, till exempel textanalys. De har dels produktifierat områden men arbetar även i forskningsprojekt såsom sentimentanalys av Twitterflöden och andra stora textmassor.

Lidberg anser att efterfrågan på verktygen finns på grund av den enorma datatillväxten. Han menar att datatillväxten blivit noterad av hela branschen och att alla stora undersökningar pekar åt samma håll. När det gäller vilka kunder som efterfrågar tjänsterna berättar Lidberg att det inte enbart rör sig om större företag utan att man befinner sig precis i början på resan för den här teknologin. Det handlar snarare om mängden data som företaget hanterar och hur mogna de är inom IT snarare än dess storlek i termer av omsättning och anställda.

SAS Institutes verktyg behandlar främst scanning och analys av stora datamängder. Deras lösning för textanalys har funnits på marknaden i ungefär tio år. Ett exempel på användande som Christopher Broxe ger är organisationer som har stora mängder PDF- eller worddokument som skall hanteras, det kan röra sig om till exempel ekonomisk rapportering eller forskningsrapporter. En stor användare av den här typen av verktyg är läkemedelsbolag. Det skrivs hundratals avhandlingar varje dag och om företaget snabbt kan hänga med och hitta det relevanta innehållet i dessa har man vunnit en stor strategisk fördel. Rent praktiskt förklarar Broxe att det fungerar så att man kopplar upp sig mot databaser med upp emot 50 miljoner publikationer vilka scannas under kort tid vilket inte hade varit möjligt utan rätt verktyg. En annan vanlig kund är banker och callcenters som vill kunna ta emot synpunkter via webbförmulär och med hjälp av dessa analysera innehåll och hitta samband i de fall kunder är missnöjda. Det kan även röra sig om att finna samband man inte är medveten om idag. Genom att spara data kan man gå tillbaka och finna vad man kallar "emerging issues" genom till exempel historisk data från servicetekniker vilken kan analyseras för att hitta problematiska samband. Hos SAS Institute ser Broxe att efterfrågan för

deras verktyg och tjänster inom Big Data har ökat kraftigt de senaste åren. Inom vissa områden och länder har ökningen varit över 100-200%. Detta tror han dels har att göra med att internet har blivit allt viktigare men också att förtroende för lösningarna och insikt hos kunderna har ökat.

Konsulter

Fredrik Kjellberg på Acando utforskar området ifall man skulle få en kund med krav på analys av datamängder i storleksordningen några tiotal terabyte eller mer. Den mängden data är svår att hantera i en relationsdatabas med bra prestanda. Här följer man vilka verktyg som kan analysera informationen inom en realistiskt tidsram. Kjellberg lyfter fram Apache Hadoop som är ett MapReduce ramverk som en lösning vilket ger möjlighet till linjär uppskalning av kapaciteten. Han beskriver hur många av verktygen som han just nu studerar har sitt ursprung hos stora "drakar" (Twitter, LinkedIn, Facebook, Yahoo) och från början är skapade för deras interna behov. De har sedan delat med sig av tekniken gratis i open source-communityn för att få draghjälp i utvecklingen.

Johan Jerresand på Affecto bedömer att möjligheten till analys av ostrukturerad information är viktigast inom Big Data. Sensordata från smarta maskiner blir allt viktigare men samtidigt produceras ännu större mängder ostrukturerad information, särskilt om man väger in tal och video. Han bedömer att det största hotet mot Big Data är att företagen inte tar språnget och lär sig den nya tekniken och de nya arbetssätten de kräver.

Användare

Mats-Olov Eriksson berättar hur King är helt datadrivna i sin verksamhet och han beskriver hur man nästan inte använder något annat än det som finns loggat. Intäkter är ett exempel på något som mäts genom att titta i datan och vidarebefordras sedan till finans som behöver det för bokslut. Produktägarna vill veta hur det går för spelarna och hur popularitet varierar, alltså uppföljning av produkterna. Eriksson menar att deras Big Data-problematik började på riktigt när de kom igång med social gaming. Genom övergången fanns det ett annat behov, verksamheten skalades upp och de gjorde en översyn där man ritade upp hur man tänkt sig dataflödet från noll. King insåg att man hade för stora mängder spelare för att man skulle kunna köra vanliga

SQL-lösningar i sin produktionsdatabas. Efter att ha testat olika lösningar bland annat in-memory-analys fann de en lösning. Eriksson beskriver hur de istället skapade ett minimalt dataobjekt till själva spelandet och loggade allt i en textfil i varje webbservar. I dessa loggar sparas händelser i rader med ett antal kolumner till exempel vilken produkt, användartransaktion och dylikt. Sedan kan varje rad i loggen urskiljas och användas. Eriksson menar att de på vis skapade en väsensskild databas jämfört med SQL.

Loggarna tas sedan i sitt rena format och läggs in i Hadoops styrsystem. På detta har de en installation av verktyget Hive i vilket man kan ställa queries. De tre första raderna i loggarna berättar för Hive hur datan skall tolkas. Detta kommer ur Hives perspektiv få det att se ut som en konventionell databas trots att det inte en indexerad databas utan enskilda textfiler som kan flyttas runt i och med att man berättat för Hive hur de skall läsas. Att lägga in en ny entry i databasen är lika enkelt som att lägga in en ny textfil i en mapp. Detta skapar en dynamisk och skalbar lösning. En av de stora fördelarna med tekniken är att man inte behöver använda dyr och avancerad hårdvara för databashantering och analys utan kan förlita sig på billig standardhårdvara vilken snabbt kan bytas ut om något går fel. Eriksson beskriver hur jobben distribueras på ett flertal maskiner med en replikeringsfaktor på tre. Detta betyder att datan finns på tre ställen och sedan samlas den i en reducer därav benämningen MapReduce.

Eriksson påpekar att MapReduce, Hadoop och framförallt Hive ännu inte är mogna tekniker och det finns mycket kvar att göra innan de är redo för stordrift där man kan lita på det i livsviktiga tillämpningar. Han menar att det för King inte är katastrof om man tappar bort en timmes data. Eriksson anser att Hortonworks och Cloudera bidrar väldigt mycket till projektet men att det troligtvis är ett par år kvar innan lösningarna är mogna. Då språket är likt SQL förväntade han sig i början samma saker men det är långt ifrån lika moget. Trots detta ser Eriksson näst intill oändligt med användningsområden för området i form av nya affärsmodeller och kostnadsbesparingar. Eriksson anser att det finns mycket att göra inom områden såsom logistik och annan affärsverksamhet där loggning av stora mängder data är möjlig.

4.1.4 Retorik och behov

Leverantörer

Microsofts Big Data-satsning beskrivs av Simon Lidberg som bred där man vill tillgodose alla tänkbara behov inom området. Han menar att branschen är i ett såpass tidigt skede att man ännu inte säkert kan veta var detta kommer sluta. Detta innebär att leverantörer måste gå ganska brett i ett initialt skede. Man utgår från olika industriprocesser i forskningsprojekt och tittar på hur man kan tillgodose rådande och kommande behov inom Big Data. Detta arbete sker både på eget initiativ och tillsammans med företag för att se vilka behov som finns. På många områden menar Lidberg att det dock än så länge rör sig om ren forskning. Lidberg påminner dessutom om att Microsoft själva är stora användare av Big Data vilket gör att det är väl insatta i problematiken. Detta kan exemplifieras med deras egen sökmotor Bing som under 2010 scannade ungefär 7 petabyte data per månad och Microsoft erbjuder även molntjänster för företag där de kör runt en miljard CPU:er.

För att hämta information om marknaden använder sig Microsoft av flera parter. Lidberg nämner IDC, Forrester och Gartner som inflytelserika aktörer inom branschen i stort. När det gäller just Big Data tillägger han även Monach research som har funnit på området länge. Den här typen av information hämtar de främst för att följa och förstå hur marknaden fungerar och vilken inriktning den tar.

Christopher Broxe på SAS Institute berättar att behov för textanalys förekommer inom alla typer av branscher i stort men att det i allmänhet rör sig om lite större företag och myndigheter. Han tror att behovet av lösningarna ursprungligen fötts ute hos kunderna och att användandet mognat mer och mer för varje år. Stora användare inom området är forskning i största allmänhet och i det området rör det sig mycket om läkemedel, medicin och sjukvård men även banker. Gemensamt för kunderna är att de har tillgång till stora mängder dokument eller information som behöver gås igenom.

Broxe är ingen stor användare av analytikerrapporter och liknande men han beskriver hur det finns andra på SAS som följer Gartner, Forrester och liknande närmare. Han är dock medveten om deras position och anser att det som sägs kring den typen av

aktörer och hur de driver utvecklingen generellt stämmer ganska väl. Han tror dock inte att SAS påverkas så mycket av detta då det är ett företag med stark ägarstyrning. Än idag leds det av samma person som grundade det på 70-talet. Han håller dock med om att den typen av rapporter generellt får stort genomslag då de läses av många intressenter, såväl kunder och uppköpare som konsulter och mjukvaruföretag.

Konsulter

Acando är ett management- och IT-konsultföretag som i nära samarbete med sina kunder identifierar och genomför bestående verksamhetsförbättringar genom informationsteknik. Managementkonsulterna inom BI har verksamhetskompetens och hjälper bolagen att strukturera arbete med att specificera behov och välja lösningar. IT-konsulterna inom BI designar och implementerar de olika BI-leverantörernas lösningar. Fokus för Fredrik Kjellbergs del har legat i att hitta verktyg och lösningar som kan hantera analysen. Många av verktygen som Fredrik just nu tittar på är open source-lösningar, men det finns även andra aktörer som paketerar lösningarna genom att göra de mer användarvänliga, lättare att förstå sig på och erbjuda support.

Kjellberg använder sig inte personligen av information från de stora analytikerfirmorna i någon större mening, men annan personal på BI läser deras rapporter. Fredrik har inte så bra tillgång till analytikers rapporter och hans ingång i området är dessutom mer ”hands-on” som systemarkitekt. Han menar dock att rapporterna spelar in på hur branschen formas. *“Det finns väl anledningar till att Microsoft har plockat upp Hadoop till exempel, de läser väl också Gartnerrapporter”*.

Affecto prenumererar på Gartners material och är dessutom en partner till firman. Det är enligt Jerresand helt rimligt att företag tar hjälp av researchföretag istället för att bedriva egen forskning. Eftersom kunderna oftast har en annan huvudkompetens är IT finns det risk för att de mest centrala delarna av hur ny teknik kan påverka kunderna riskerar att försvinna i bruset. Det kan hända att man tar med sig Gartners information ut till kunder för att referera till när man presenterar lösningar och tekniker då han ser det som mycket viktigt att förklara för kunder vad förutsägelser om ny teknik och förändringar kan innebära för deras del. Han anser att Gartners “predictions” är särskilt bra för att visa detta för kunderna då de illustrerar kommande förändringar genom påståenden som kan vägleda tankar kring utvecklingen och inte behöver, eller

ska, bokstavstolkas. Rörande Gartners inflytande inom Affectos verksamhetsområde säger Jerresand:

“Det finns många som är skeptiska till de här researchinstituten, men det är som jag brukar säga att det de säger blir förr eller senare en sanning oavsett om man vill det eller inte. De hittar ju inte bara på utan de fångar tidigt upp signaler för vad kunderna har behov av, vad kunder gör.”

“Det finns kunder som är skeptiska till researchföretagens information, men då tycker jag att man är lite fel ute, för man hittar ju inte på saker utan det är kopplat till en behovsbild som finns.”

Jerresand har märkt att kunderna vill vara tidigt ute med ny teknologi. Han tror att man som anställd med uppgift att studera nya teknologier har att välja mellan att helt bilda sig en egen uppfattning eller så söker man vägledning i material från researchföretag och avgör om lösningarna passar den egna verksamheten. Han tror att man fattar samma, eller liknande, beslut som sina konkurrenter vilket kan göra att självuppfyllelsen av analytikernas påståenden snabbas på.

Jerresand anser att Affecto befinner sig mellan utbudssidan och efterfrågesidan, men förståelsen för kundens behov kommer alltid i första hand. Beroende på om man kommer in tidigt eller sent i ett projekt påverkar var man hamnar mellan utbuds- och efterfrågesidan. Om man kommer in tidigt är man oftast djupare involverad i processen och har större chans att agera med utgångspunkt från efterfrågesidan medan man oftast endast får presentera en lösning från en leverantör om man kommer in sent.

Användare

Mats-Olov Eriksson berättar hur King ser till sina egna behov, de lägger ingen stor vikt vid vad analytikerfirmor säger. Man bevakar omvärlden, men det går ganska bra att göra i det intresseforum som finns och dessutom ser man vad som händer runt omkring sig. Eriksson menar även att analytiker och tyckare är så pass “corporate” att man på King inte riktigt kan relatera till deras rapporter. Han menar att man har fullt upp med sin egen data och den egna BI:n handlar bara om att agera översättare av

data till analyserbar information och används inte för omvärldsinformation. Men det kan finnas ett affärsmässigt värde att tala om.

Han menar att de flesta stora aktörerna i vissa fall i princip kan fungera som lobbyister. Eriksson menar att de har både inflytandet och musklerna för att inte bara hitta en förändring utan i princip skapa en. Eriksson berättar hur Big Data tidigare växt i någon form av undergroundrörelse men de som kan paketera det snyggt av de stora jättarna som Oracle, Microsoft och IBM sedan kommer att kunna prata om det här på ett sätt som kommer få till exempel NASDAQ att lyssna. Eriksson tillägger att han gärna vill ta bort lite av hypen från Big Data. Han anser att det är viktigt att lyfta blicken lite och fundera på vad man egentligen vill ha ut som användare. Eriksson menar att MapReduce-teknik inte är något som skall användas förrän behovet verkligen finns. Han har träffat många företag som fastnat i Big Data men det enda de behöver enligt honom är en duktig systemutvecklare som bygger en bra databas för det de vill ha är analys på. Trots detta menar Eriksson att deras Big Data-lösning överträffat deras förväntningar då det möjliggör att man är kreativ med det. Datan är orörd och man kan berätta för systemet hur den skall användas.

Han anser att det går att komma väldigt långt på området om man gör saker själv, bara man ser till att ha rätt folk inhouse. Eriksson menar att alla företag som har runt 100 anställda nästan alltid tjänar på att ha det just inhouse då de kan tjäna hur mycket som helst på det som kan uträttas med teknisk kompetens och verksamhetsförståelse.

“King har aldrig varit något för en hyllösning. Saknar vi något bygger vi det. Men det hade säkert gått att lösa med någon form av leverantör och en triljard konsulttimmar.

4.1.5 Sammanställning av intervjuer

	Leverantör	Konsulter	Användare
Vad är Big Data?	3V Stora mängder strukturerat/ostrukturerat. Problem är storlek, samkörning & antal händelser.	Ostrukturerade datamängder >1 terabyte, tidskrävande analys 3V+komplexitet	För stort antal transaktioner för konventionell lösning
Verktyg & teknik	Plattform för stora strukturerade datamängder, verktyg (Hadoop) för analys av ostrukturerad information Scanning och analys av stora datamängder	Verktyg för snabb analys av stora datamängder ex Hadoop Verktyg för analys av ostrukturerad information och sensordata	Hadoop och Hive: Hantering och analys av loggfiler
Retorik & behov	Kraftig datatillväxt bör utnyttjas, affärsmöjligheter Hitta strategiska fördelar, kan leda till effektivisering, finna samband	Samla information, användning behöver ej vara klar Affärsmöjligheter, konkurrensfördelar/existens	Endast eget behov skraddarsydda lösningar för att lösa en problematik som sedan gett synergieffekter
Vem styr utveckling?	Behov hos de som hanterar stora datamängder. Utveckling startade hos Google, Yahoo. Användares behov först	Initialt open source, drivs nu av analytikerhype som lett in leverantörer på området. Analytiker ser framtida behov, många hakar på, symbios.	Mikro: Själva. Problem uppstår, lösning tas fram. Makro: Analytiker & leverantörer som riktar in sig på större företag
Hur stödjer det ekonomistyrning?	Användning inte helt klar, utökar analysmöjligheter. Bättre beslut och effektivisering. Textanalys stödjer uppföljning/utv. av handlingsstyrning.	Mer relevant analys, se trender och samband. Beslutsfattande, utveckla affärsmodell (segmentering/differentiering)	Allt, helt datadrivna, tom finansdata tas från loggar, produktutveckling, tidigt fånga upp avvikelser

Tabell 2. Sammanställning av empiri

Källa: Egen

5. Analys

I analysen jämför vi vår empiri med referensramen inom följande områden: Begrepp, relation till BI, användning, Big Data i ekonomistyrningen och avslutas med att diskutera Big Data som ekonomistyrningsinnovation.

5.1 Begreppet Big Data

Vi har under arbetets gång märkt att begreppet Big Data än så länge inte har fått en enhetlig definition. Konsultfirman McKinsey (McKinsey Global Institute, 2011) väljer att endast fokusera på volymperspektivet där huvudproblemet är att de gamla databasstrukturerna inte kan hantera alltför stora volymer data på ett tillfredsställande sätt. Andra aktörer (Russom, 2011) (Zikopoulos, et al. 2012) väljer en bredare syn där variation och hastighet på data läggs till beskrivningen. Det förekommer även ibland ett fjärde perspektiv (Rhodin, 2012) voracity, som ska spegla den ökade efterfrågan på information från användare. De flesta respondenterna nämner volymproblematiken som viktig, men även antalet transaktioner som kan definieras som hastigheten på skapandet av data. Flera respondenter är inlästa på det etablerade konceptet "3V" som används i den retorik stora aktörer på marknader använder. Det finns åsikter om att en fjärde komponent bör vägas in (Jerresand, Lidberg) men det verkar vara oklart exakt vad den ska spegla. Jerresand pekar på komplexiteten som enligt honom ökar kraftigt på grund av de tre andra komponenterna.

Det tycks finnas en allmänt utbredd insikt att de växande datamängderna är ett problem som måste hanteras i de verksamheter som berörs. De tre perspektiv vi intervjuat är i huvudsak med på upplägget av att formulera Big Datas huvudområde som 3V, medan användaren framförallt talar mer utifrån sitt eget perspektiv. Det är troligt att tekniken och de erbjudna lösningarna inte fått någon enhetlig definition på grund av att området Big Data fortfarande är i en utvecklingsfas. Vi anser att de tre V:na, volume, variety och velocity, på ett bra sätt beskriver utmaningarna som den moderna IT-användningen medför. De stora ostrukturerade datamängderna hos företag växer och måste dels lagras men många är även intresserade av att använda dem för att göra analyser. Lagring och beräkningskraft är idag inte på samma sätt som tidigare en kostnadsfråga då kapaciteten har ökat samtidigt som de relativa priserna har sjunkit. Det är därför realistiskt för organisationer att besluta sig för att börja lagra

och använda sig av data på ett annat sätt än tidigare.

5.2 Relation till BI

Traditionella analysverktyg eller beslutsstöd är enligt Baars & Kemper (2008) uppbyggda i tre olika lager, datalager, tillgänglighetslager och presentationslager. Hittills har denna struktur fungerat tillfredsställande för att hantera de volymer och sammansättningar av data som förekommit. Det finns dock aktörer som menar att de traditionella BI-lösningarna inte längre räcker till för att hantera den ökade volymen och komplexiteten i moderna organisationer där den ostrukturerade informationen enligt Shilakes & Tylman (1998) och TDWI (Russom, 2007) står för mellan 50-80% av organisationers samlade information. Ostrukturerad data kräver nya tekniker då det inte gått att hantera i stora mängder med traditionell BI-teknik (Russom, 2007).

TDWI (Russom, 2011) påpekar även att insamlings- och rensningsfasen riskerar att sortera bort värden som de tar för felaktiga men i själva verket kan vara viktiga för att nå nya insikter. En undersökning, TDWI (Russom, 2011) har också visat att användare upplevde att deras lösningar inte kunde hantera tillräckligt stora datamängder samt att laddningen av data var för långsam. Vår intervju med Mats-Olov på King stödjer dessa teorier då de tidigare använt traditionella BI-lösningar men inte varit nöjda med prestanda och flexibilitet. I de övriga exemplen rör det sig om hantering av ostrukturerad data från olika kund-, sensor-, och externa källor. När det rör textanalys är det inte enbart ett prestandaproblem utan funktionaliteten ligger utanför de traditionella BI-lösningarna och kräver analys av stora mängder information.

Big Data skulle kunna ses som en vidareutveckling på BI-lösningar då kunder idag vill göra mer analys på nya sätt (Prentice, 2011). Sett ur detta perspektiv skulle man kunna placera in Big Data-verktyg och dess lösningar i data/accesslayer enligt Baars & Kempers (2008) modell. Den huvudsakliga tekniska påverkan från Big Data har varit i dessa lager, till exempel genom distribution av datorkraft med hjälp av parallella processer, men det är möjligt att nya typer av analyspresentation blir nödvändiga i framtiden.

Skriletz (2002) menar att företag bör anstränga sig för att använda sin BI strategiskt och inte endast stödja och effektivisera det operativa arbetet vilket kan begränsa värdet som BI kan bidra med. Den enda begränsningen för BI som finns enligt författaren är möjligheten att samla användningsbar data för att senare använda för analys. Big Data-verktyg ökar markant möjligheterna för att använda olika typer av data i analysarbetet. Med rätt kompetens kan man enligt författaren använda alla typer av output för att bättre förstå sin verksamhet och söka nya möjligheter. Samtliga respondenter berör detta när de talar om möjligheterna att hantera och analysera datamängder som tidigare inte var möjligt och hur man ska driva detta till strategisk nytta. Det kan tänkas att utformning av analysarbetet kan komma att behöva ändras när man går från att analysera information som man på förhand klassat som relevant till att synen breddas till att all information kan vara intressant när den sätts i rätt samband.

5.3 Användning

5.3.1 Teknik

Det tycks som att en del av Big Data-hypen ligger i att informationsmängden håller på att springa ifrån den databasteknik baserad på relationsdatabaser där data hanteras och hämtas via SQL vilken varit etablerad sedan 1970-talet (IBM, 2012). För verksamheter som hanterar information så att något av de tre V-kriterierna realiserar uppstår en problematik. Tekniker såsom Fredrik Kjellberg och Mats-Olov på King ser MapReduce och NoSQL som den bästa lösningen på problematiken. Inom området är det öppna mjukvaruramverket Hadoop väl etablerat och ansett. Hadoop ses av Gartner i rapporten *Top 10 technology trends impacting information infrastructure 2012* (2011) som en av de tio teknologitrender som kommer att få störst genomslag under 2012. Ramverket består av ett filsystem samt ett tillhörande verktyg för parallell data processing, MapReduce. Lösningen möjliggör hantering och analys av data i flera processer på flera maskiner parallellt och ställer inga krav på att informationen finns in en relationsdatabas. Även hårdvaruutvecklingen är på väg åt samma håll med fler kärnor i processorerna och möjligheten att koppla ihop flera servrar vilket gör tekniken skalbar.

Leverantörer av mjukvarulösningar såsom Microsoft, IBM, Cloudera och Amazon har uppfattat intresset kring Hadoop och har, eller är i färd med att, lansera egna paketeringar vilka i grunden bygger på och bidrar till ramverket. I flera av de exempel på användning vi beskrivit förekommer Hadoop i form av leverantörspaketerade lösningar som den från IBM som Vestas använder sig av vid sin analys av ostrukturerad information. Även King använder sig av Hadoop, men till skillnad från andra exempel har de utifrån open source-communityn och en extern fristående konsult utformat sin egen lösning. Mats-Olov Eriksson berättar hur de efter en översyn fann detta som en bra lösning för dem. Han påpekar dock att Hadoop och dess verktyg är långt ifrån mogna som produkter och att de kräver tillåtelse av en felmarginal och ställer höga krav på systemutvecklare och teknisk personal för att implementera. Hadoop-baserade lösningar behöver absolut inte vara någon universallösning för analys av betydande datamängder men det är idag det mest framstående alternativet när en viss komplexitet uppnåtts. Denna bild stärks av Forrester Wave som beskriver hur marknaden idag är full av omogna Hadoop-paketeringar även på enterprisenivå. Det lär däremot dröja ett antal år innan det är färdigt för användning inom branscher där verksamheterna innefattar hög risk och felaktiga beslut kan få allvarliga konsekvenser (Eriksson, 2012).

5.3.2 Praktik utifrån casen

Möjligheten att snabbt kunna analysera stora datamängder kan enligt vår studie ge användare fördelar på flera områden. Station Casinos har till exempel fått en bättre förståelse för sina kunder genom att logga stora mängder information från en rad olika källor och sedan samköra denna för att skapa förståelse för kundbeteenden. Detta har lett till att man har ökat sina kundsegment från 40 st till 160 st. I denna tillämpning är man som företag sannolikt ute efter att ge bättre service till sina kunder vilket troligtvis leder till bättre kundnöjdhet, vilket ofta kan föregå ökade intäkter.

Även för King har deras användande av Big Data gjort att de kan förstå sina kunder bättre. Det är mycket viktigt att snabbt kunna följa upp hur deras kunder reagerar på förändringar eller nya varianter av spel för att inte riskera att förlora spelare. Genom att göra sambandsanalyser kan de upptäcka orsaker till kundbortfall och se samband relaterat till svårighetsgrad och prestanda i spelen. De kan då snabbt agera och ändra

spelen efter användarnas villkor. Numera hämtar man i stort sett all information som används för analys genom lösningen och eftersom all information lagras kan de analysera väldigt flexibelt, utan att i förväg behöva definiera hur insamlingen ska ske. Detta kan mycket väl vara av stor vikt som strategisk konkurrensfördel. Genom snabba analyser har man möjligheten att kunna se hur deras spel tas emot av spelarna och fokus är att hålla en hög kundnöjdhet och second day retention (andel spelare som återvänder).

AFA försäkring har dels effektiviserat sin insamlingsprocess för skadehantering genom att kunna överge den manuella granskningen av skaderapporter samt den kodning av innehållet som tidigare var en förutsättning för att kunna göra analys. Genom detta har kostnaderna för arbetstid minskat samtidigt som analyserna sannolikt blivit mer träffsäkra då risk för att missa viktiga faktorer har minskat. Enligt Broxe har AFA lyckats få helt nya insikter om faktorer som tidigare sannolikt fallit utanför intresseområdet men som nu kunnat fångas upp genom att textanalysen ser sambanden. För AFAs del innebar alltså användningen av textanalys fler vinster än de som var skälet till investeringen från början.

För Vestas del handlar Big Data om att samköra stora mängder data av olika typer bland annat för att kunna fatta bättre beslut om olika lokaliseringar av vindkraft kommer att vara lönsamma baserat på väderdata, sensordata, GPS-information och dylikt. Man arbetar med stora komplexa investeringar som kräver noggrann analys för att man ska kunna avgöra lönsamheten. Den noggranna insamlingen och analysen gör även att man kan anpassa underhåll och service till tidpunkter då behovet är som störst, men innan kraftverken går sönder, samt då kraftproduktionen spås vara låg.

Christopher Broxe ser stor möjlighet till användning av Big Data i form av textanalys inom exempelvis läkemedelsforskning där läkemedelsföretag har ett behov av att få reda på innehållet i stora mängder forskningsrapporter, eller hela databaser, som kan relateras till egna forskningsområden. Det hade inte varit möjligt att få en överblick utan att använda sig av textanalys på dessa områden. Lidberg möter ofta på marknadsförare som vill få reda på hur deras företag och produkter framställs på nätet och i sociala medier genom sentimentanalys. Dessutom nämner han tillsammans med både Jerresand och Kjellberg möjligheterna för fordonsindustrin att samla in data från

fordon och använda det för att till exempel se när det finns behov av underhåll eller tolka mönster av användningen.

När det gäller hur lösningarna ska användas upplever vi i vår studie att det finns en skillnad mellan utbuds- och efterfrågesidan. Leverantörer och konsulter på efterfrågesidan talar om att man bör spara all data utan att nödvändigtvis för tillfället ha ett klart användningsområde för den. Efterfrågesidan tenderar snarare att söka en lösning när det uppkommer ett klart behov.

5.4 Big Data i ekonomistyrningen

Malmi & Brown (2008) definierar ekonomistyrning i termer av olika koncept vilka sorterar de kontrollmekanismer som skall styra organisationers verksamhet i rätt riktning. Ytterst handlar ekonomistyrningen om att stödja företagets arbete med att forma och uppnå dess mål. Begreppet i sig innefattar planering, genomförande, uppföljning och anpassning av verksamheten (Ax, Johansson, & Kullvén, 2009). För att fatta rätt beslut kring dessa områden krävs det att man har tillgång till ett så bra underlag som möjligt. Detta för att analysera trender och planera verksamheten utefter de förutsättningar som råder.

Enligt Samuelson (2001) syftar det han kallar formella styrmedel till att fastställa och följa upp den strategi som utarbetats från affärsidén. Detta passar in på informationshantering och analys, och här skulle användande av Big Data kunna stödja uppföljning och härleda händelser som påverkar värdeskapande aktiviteter. Även Simons (2000) menar att mätning syftar till att realisera strategin genom att kvantifiera mål.

I takt med att informationsmängderna ökar finns mer och mer information tillgänglig för analys. Porter (1985) menar att ökad tillgång till information och möjligheten att hantera och överföra denna möjliggör konkurrensfördelar och nya affärsmodeller. Detta i så stor mån att det kan leda till skiften mellan konkurrenter inom branscher. Jerresand menar att flera industrier står inför en revolution i och med Big Datas framväxt. Enligt honom handlar det på kort sikt om konkurrensfördelar medan det på längre sikt handlar om överlevnad. Ett exempel är möjligheten till att kunna

segmentera sitt kunderbjudande på ett mer detaljerat sätt, vilket kan göra att man kan erbjuda produkter som är bättre anpassade till olika kundgrupper.

De exempel vi har hittat visar på att företag försöker använda så mycket av den internt genererade informationen som möjligt kombinerat med extern information för att skapa sig en heltäckande bild av sin egen verksamhet samt omvärlden och till exempel kunna anpassa sig efter de rådande situationer som kan påverka dess verksamhet eller de kunder man vill tillfredsställa. Jerresand nämnde att de nya elmätarna som rapporterar elförbrukning mer frekvent än tidigare skulle kunna påverka elhandelsbolagens affärsmodell och möjliggöra för dem att segmentera sitt erbjudande i större utsträckning.

Detta bygger på att man utför mätningar och sedan analyserar dessa, något som passar in på Malmi & Browns egenskaper för cybernetik. Mer detaljerad information gör det möjligt att kvantifiera fler fenomen och på så vis göra variansanalys och hitta källan till problemet. Genom att använda mer mätning och förhoppningsvis göra bättre analyser och förstå orsaker till situationer kan man nå ännu närmare sina mål. Här passar Kings verksamhet väl in då deras produkter är i sin helhet mjukvara vilket möjliggör för dem att mäta precis allt som sker, därför är en bra utformad cybernetik centralt för dem. Genom att de använder loggfiler för i stort all analys behöver inte använda olika sätt och olika input för att skifta mellan analys av finansiella och icke-finansiella nyckeltal utan kan strikt fokusera på det som anses vara kritiska faktorer och ger den bästa bilden av verksamheten. I deras fall suddas alltså gränserna ut mellan de olika typerna av nyckeltal och fokus hamnar istället på möjliga användningsområden snarare än var informationen kommer ifrån.

Sentiment- och textanalys kan även användas för att sammanställa feedback för ett bredare analysmaterial. Att rationalisera bort risken för olika tolkningar inom regler och procedurer minimerar risken att fall behandlas olika. Det finns också möjligheter att effektivisera de policys och procedurer som Malmi & Brown nämner. AFA-försäkrings användning av textanalys har enligt Broxe effektiviserat företagets administrativa arbete med skaderapporter då de automatiserat tidigare manuellt arbete. Man kan se det som att man minskat risken för fel i procedurerna genom att man numera använder textanalys som inte riskerar att missa viktiga samband.

Utvecklingen av ekonomistyrning har genom åren gått från enbart finansiella mått till att innefatta även icke finansiella mått som kundnöjdhet, omvärldsinformation och relationer vilket i förlängningen kan bidra till uppfyllnad av finansiella mål. Insamling och analys av den information som finns kring dessa områden kan leda till en ökad förståelse för de områden som bidrar och skapar värde. Den uppföljning som Station Casinos gör på sina kunder gör att de kan finna hur de med enkla medel kan öka kundnöjdheten och fokusera på vad som skapar värde ur ett kundperspektiv. King kan genom sin analys av loggfiler se relationer mellan kundernas beteenden och de förändringar i spelen som företaget gör. På så sätt kan de avgöra vilka åtgärder som skapar mest kundvärde. I fallet med Vestas gör omvärldsinformationen att de kan minimera risk och öka chansen till måluppfyllelse för ett investeringsprojekt eller lokalisering.

Man bör dock notera att Skrietz modell över operativ och strategisk användning av BI inte passar helt perfekt in i Malmi & Browns paket av ekonomistyrningmetoder då den syftar till att söka helt nya frågor och nya områden att analysera. Cybernetik i Malmi & Browns modell rör först och främst att övervaka redan identifierade områden och frågor vilket kan innebära att användning av Big Data faller utanför modellens avgränsningar. Det är möjligt att framtida ekonomistyrningsmodeller kan inkludera ett större fokus på icke-finansiella nyckeltal och ett perspektiv där sökande av, och strategisk hantering av, ny kunskap tas i beaktande för att spegla de möjligheter som skapas vid användning av nya typer av analys.

5.4.1 Big Data lösningar som ekonomistyrningsinnovation

Abrahamson beskriver i sin Management-Fashion-Setting Process (1996) hur processen för managementmodens fastställande och spridning ser ut. Denna modell kan ur vårt perspektiv appliceras på processen för vårt område med Big Data-lösningar som en ekonomistyrningsinnovation. På utbudssidan agerar gurus och analytiker från de stora analytikerfirmorna. De conceptualiserar de trender och behov de ser inom området och formar på så sätt den retorik och de designelement som kommer etableras. Enligt Abrahamson (1996) skapar man genom tekno-ekonomiska krafter en medvetenhet hos användare om att ny teknologi kan höja den potentiella

prestationen över den existerande. Jerresand har märkt av att kunder blivit mer intresserade av nya teknologier, vilket kan vara ett tecken på att tekno-ekonomiska krafter påverkat hur Affectos kunder tänker kring ny teknologi. På detta vis får användarna sina efterfrågade lösningar namngivna och uppmärksammade i analytikernas rapporter. Rapporterna läses av potentiella användare på efterfrågesidan och implementeringskonsulter. Konsulternas roll beror enligt Johan Jerresand på när de kommer in i processen. Han menar att Affecto rör sig mellan utbudssidan och efterfrågesidan, men förståelsen för kundens behov kommer alltid i första hand. När utbudssidan etablerat sin retorik och de designelement som kretsar kring denna utifrån de behov som finns på efterfrågesidan tar leverantörerna till sig dessa och paketerar dem för att fånga upp det behov som finns. Flera respondenter beskriver just detta fenomen och hur analytikerfirmor som Gartner, Forrester och IDC är bra på att fånga upp intressen och behov på marknaden.

I fallet med Big Data och de open source-projekt som ligger till grund för många lösningar finns dock en skillnad gentemot Abrahamssons modell då teknologin för att hantera de stora datamängderna från början skapats av aktörer i det privata näringslivet såsom Google, Amazon och Yahoo. De hade själva behov av att analysera stora mängder data och utvecklade egen teknik för att sedan släppa den fri och på så vis få hjälp med systemutveckling. Däremot har tekniken än så länge inte nått någon bredare adoption, något som enligt Eriksson kan bero på att fokus hittills inte har legat på användarvänlighet utan snarare på utveckling och prestanda. Det är dock här som utbudssidan kommer in på allvar och det verkar vara där vi befinner oss idag. Nu arbetar flera stora leverantörer som med att paketera och göra tekniken mer användarvänlig för att kunna erbjuda den till en bredare skara användare utan kompetens i systemutveckling. Mats-Olov Eriksson på King beskriver att kompetensen krävs för att på ett framgångsrikt sätt utveckla och implementera en Hadoop-baserad MapReduce-lösning på egen hand.

Det rimligt att tro att många företag hamnar i takt användningen av Big Data ökar. Det ökade intresset som märkts under det senaste året kommer troligtvis göra att företagen vill försäkra sig om att inte hamna efter sina konkurrenter vilket gör att de adopterar teknologin även om de inte är helt mogna för det. I vår empiri sticker användaren King ut, de verkar ha gjort ett efficient choice när det gäller användningen

av Big Data. Man ställdes inför det faktum att existerande lösningar inte kunde leverera den prestanda och flexibilitet man krävde och kunde därför inte tillgodose företagets krav. Genom sin egna höga, men även delvis inhyrda, kompetens och nytänk har man skapat en egen lösning som efter cirka nio månader börjar ta form på riktigt enligt Mats-Olov Eriksson. I dagsläget är forced selection-perpektivet än så länge inte aktuellt för Big Data men i framtiden skulle det vara möjligt att krav från kunder, leverantörer och andra intressenter leder till att företag behöver hantera och analysera sina datamängder på ett mer avancerat sätt än idag.

5.4.1.1 Designelement

Med designelement menar Ax & Bjørnenak (2011) den tekniska specifikationen på en innovation. Gartner definierar i sin Hype cycle en rad tekniska områden som kan kopplas till Big Data. Som branschauktoritet och trendsättare på utbudssidan enligt Abrahamssons Management fashion setting process namnger de innovationer som finns på agendan. I deras hype cycle över BI tar de upp Big Data-relaterade designelement som bland annat content analytics, text analytics och sentiment analytics vilka skall klara att samla in och analysera olika former av ostrukturerad information såsom text, pdf, bilder, videor och inlägg på sociala medier. Då ett av de hetaste områdena inom Big Data är hantering av stora ostrukturerade datamängder genom Hadoop är paketeringar baserade på detta givetvis ett framstående element. Det förekommer vid sidan av Hadoop verktyg som kan bidra till insamling och analys av informationen som de innehålls- och textanalysverktyg som SAS Institute och en rad andra leverantörer erbjuder.

När det gäller den här biten finns det två typer av leverantörer, dels de som paketerar och levererar helhetslösningar såsom Microsoft, IBM och SAP, och de som levererar specifika verktyg vilka enbart hanterar ett tekniskt område. Sett till den bild vi har fått av området skulle man kunna dela in de verktyg som används i Big Data-lösningar i följande grupper:

-Verktyg för analys av stora datamängder vilka i vissa fall kan stödjas av in-memory-analys för snabbare prestanda. Dessa lämpar sig bäst för de användare som har strukturerad data.

-Verktyg för analys av stora datamängder som innehåller en betydande andel ostrukturerad data. Dessa skall klara av att scanna och analysera och är ofta baserade på Hadoop. Dessa paketeras tillsammans med kringverktyg och görs användarvänliga av de stora aktörerna.

-Specifika verktyg som stödjer insamlandet och analys av informationen såsom text mining, verktyg för sentiment analys av social networks till exempel. Detta är verktyg som oftast tillverkas av specialiserade leverantörer men de stora jättarna har även de sina varianter. Dessa syftar till att adressera specifika användningsområden till exempel den textanalys AFA försäkring gör på sina skadeanmälningar.

5.4.1.2 Retorik

Ax & Bjørnenak (2011) talar om retorik som en komponent av en ekonomistyrningsinnovation. Denna används av utbudssidan för att övertyga potentiella användare om innovationens värde och hur den ska användas. Som vi tidigare fastslagit består utbudssidan av flera aktörer såsom media, gurus, leverantörer, analytiker. Just analytikerfirmor såsom Gartner, IDC och Forrester har ett stort inflytande inom IT-branschen, de bedriver forskning, rådgivning och anordnar konferenser för forskare och näringsliv. Enligt Pollock & Williams (2010) speglar dessa aktörer inte bara rådande marknadssituation utan de bidrar aktivt till att forma denna. De är tidiga med att känna av kommande trender och behov för att plocka upp och konceptualisera detta till framväxande teknologier. Genom konceptualiseringen sätter de både områden och verktyg på kartan hos både användare och leverantörer.

Användarna ser möjliga lösningar och leverantörerna ser vad användarna kommer att efterfråga och paketerar detta. Detta samspel beskrivs av Jerresand på Affecto som menar att utvecklingen på marknaden sker som en symbios och inte behöver innebära något negativt. Han anser att analytikerna ser till en behovsbild som de bedömer kan uppstå i och med att ny teknik påverkar företagens miljö och att ser det som rimligt att företag vänder sig till experter då de sällan har sin kärnkompetens inom nya former av IT. De har antingen att välja på att helt och hållet skapa sig en egen bild alternativt att ta hjälp av hur researchföretag ser på frågan.

Det finns en viss skillnad i retorik mellan användar- och utbudssidan. Utbudssidan talar om hur användare måste ta till vara på datatillväxten för att kunna effektivisera och ta vara på nya affärsmöjligheter med hjälp av deras paketerade lösningar. Samtidigt har de inte alltid bilden av de direkta fördelarna av användningen helt klar. Användarsidan talar mera om eget behov, Mats-Olov Eriksson på King menar att de inte bryr sig särskilt mycket om vad analytikerfirmor och leverantörer på området skriver i sina rapporter. De hade ett behov att finna en lösning för ett specifikt problem. Han anser att han kunnat skapa sig en bra bild av Big Data-området genom olika nätforum utifrån företagets förutsättningar.

Han menar att den bild som målas upp av utbudssidan är anpassad för en enterprisenivå vilket inte passar alla användare. För dem handlar det mera om att ha rätt kompetens och finna egna lösningar än att vända sig till någon universalpaketering. Konsultföretagens roll och retorik tycks påverkas av vilken roll de har i adaptationen av teknik, Fredrik Kjellberg på Acando menar att de verkar på både utbuds- och efterfrågesidan då det både arbetar med att specificera behov, välja och implementera lösningar. Jerresand på Affecto menar att man befinner sig på olika platser mellan utbuds- och efterfrågesida beroende på i vilket stadiet av projekten man kommer in. Ju tidigare, desto större chans att förstå kundens behov och därför vara mer på efterfrågesidan. Om man kommer in sent får man oftast bara presentera en lösning.

6. Slutsatser

Big Data är ett förhållandevis nytt fenomen som är starkt på uppåtgående i både intresse och utbredning. Teknologier som ofta associeras med området såsom Hadoop och MapReduce är än så länge inte mogna och det finns ganska få exempel på etablerade användningsområden förutom de som hanterar stora datamängder på internet. Samtidigt pågår arbetet från leverantörernas sida att paketera och tillgängliggöra tekniken för den breda massan, men även att förbättra tidigare etablerade tekniker så att även dessa klarar av att hantera problematiken. Många aktörer på utbudssidan förespråkar att möjligheterna är nästan oändliga vilket gör det svårt att bilda sig en uppfattning om fenomenet.

6.1 Vad innebär Big Data enligt leverantörer, konsulter och användare?

De olika perspektiven är i stort enade angående begreppet Big Datas innebörd och problematikens existens. Vi kan dra slutsatsen att Big Data berör teknologier för att hantera och analysera mycket stora, ofta ostrukturerade, datamängder vilka kommit som en följd av den tekniska utvecklingen och den växande aggregerade datamängden. Beskrivningen kring de tre V:na volume, variety och velocity är väl utspridd i branschen. Den ger en bra överblick av problematiken vilken lett till att företag som använder sig av mycket data har sökt nya metoder för hantering av data under det senaste decenniet både vad gäller storlek, olika typer av data samt hastigheten på skapandet av data.

Big Data kan enkelt beskrivas som ett begrepp vilket beskriver det fenomen som uppstår när de etablerade analysmetoderna inte längre klarar av att analysera en stor eller ostrukturerad datamängd på tillfredsställande tid. Det finns många starka intressen som verkar för att sprida innovationer i form av lösningar av den här typen. Big Data behandlar dock ett problem som är en del av trend som kommer i och med den tekniska utvecklingen. Självklart finns det en skillnad i retorik mellan utbuds och användarsidan. Utbudssidan talar om storskalig analys som ett måste för att på kort sikt skapa och bibehålla konkurrensfördelar och på lång sikt för att överleva medan användarna pratar om det som en lösning för ett specifikt problem vilket sedan medfört möjligheter till ytterligare fördelar.

I och med att mängden och möjligheterna till insamlande av data ökar så kraftigt kommer allt fler företag växa ur de traditionella databastekniker som utvecklades när informationsmiljön var en helt annan. I och med att föregångare på användarsidan inom hantering av stora ostrukturerade datamängder har utvecklat egna metoder och släppt dessa gratis har teknik tillgängliggjorts för att bibehålla prestanda genom att distribuera beräkningskraft på flera enheter utan krav på dyr och avancerad hårdvara. Just priser på hårdvara är något som har bidragit till att driva på utvecklingen inom Big Data då det inte längre är ett beslut av stor ekonomisk vikt att börja samla på sig mer information.

Även om de användare som besitter rätt kompetens kan utveckla sina egna skräddarsydda lösningar kommer den paketering och integrering de stora leverantörerna erbjuder att göra tekniken tillgänglig för allt fler användare på sikt. Utbudssidan är självklart intresserad av att sätta nya områden på användarnas agenda men i fallet med just Big Data är förekomsten av informationen och därmed både problematiken och möjligheterna ett faktum som alla aktörer har insett. Det är möjligt att Big Data som begrepp kommer att dö. I takt med att tekniken integreras med de stora applikationerna inom BI och analysmöjligheter samt informationstillgång ökar kommer de tillgängliggöras för alla användare utan att de behöver reflektera över problematiken.

6.2 Hur kan Big Data stödja ekonomistyrningsarbetet?

Big Data har i grunden varit ett IT-problem i och med att datamängderna i världen stigit såpass kraftigt. Ändamålet har från början inte varit att skapa en förbättrad ekonomistyrning utan det är något som har tillkommit i och med att möjligheten att hantera relevant information ökat. Till en början var det i huvudsak frågan om att bara hantera de stora datamängderna. Genom möjligheterna att inkludera stora mängder ostrukturerad information har inriktningen på analysmöjligheterna breddats avsevärt. I takt med att nya teknologier växt fram, vilket har utvecklat lagrings- och analysmöjligheterna samt IT-användning i stort, har utbudssidan bestående av i huvudsak analytiker, leverantörer och till viss del konsulter spridit dessa och vidgat vyerna för en ökad IT-användning inom ekonomistyrning. Leverantörer och intressenter inom branschen tar till sig dessa trender samt formar retorik kring

begrepp och teknologier vilka paketeras och presenteras i en form som kan liknas vid ekonomistyrningsinnovationer. Hur dessa skall appliceras i ekonomistyrningen och vilka effekter de på sikt kommer få är idag inte helt klart. Det är möjligt att de modeller vi idag har kommer fortsätta omformas och anpassas till de nya förutsättningar som uppkommer.

Fram tills nu har analysarbete på företag vanligtvis följt proceduren där man i förväg vetat vad man velat ta reda på vilket har påverkat designen av verktygen för att hantera och analysera informationen. Nu ges möjligheten att kunna analysera vilka källor som helst, både strukturerade som ostrukturerade, när som helst och dessutom spara informationen för framtida, idag okända, behov. Detta kan leda till att man hittar viktiga svar på frågor som man inte ens är medveten existerar. Att ha tillgång till rätt information kan innebära stora möjligheter om man är uppfinningsrik och kan utforska samband som gör att man kan bli medveten om saker andra aktörer på marknaden inte ser och på så sätt rikta sin strategiska ekonomistyrning.

Möjligheterna att samla och inkludera information på sätt som tidigare inte gjorts eller varit möjligt kommer troligtvis påverka många processer inom organisationer. Områden inom ekonomistyrnings cybernetik som uppföljning och mätning kan bli bättre. Det kan även få stor inverkan på områden som prognostisering och strategiska beslut då dessa kan baseras på ett betydligt mer omfattande underlag. Möjligheterna att effektivisera och automatisera analyser som tidigare gjorts manuellt är även dessa intressanta ur ett ekonomistyrningsperspektiv.

6.3 Avslutande diskussion

Avslutningsvis anser vi att området Big Data handlar mycket om förståelse för de datamängder som numera skapas och att detta kommer att bli allt viktigare ju längre utvecklingen går. Analysmöjligheterna blir väldigt mycket större och mer komplexa jämfört med traditionell Business Intelligence och ger en föräning om att strategisk informationsanvändning kommer att bli viktigare i framtiden. Det är troligt att ägandefrågor gällande data kommer att bli allt större då tillgång till information kan växa som konkurrensfaktor. Mycket kommer att handla om uppfinningsriktighet och hur man med kreativt tänkande kan finna samband att dra fördelar av. Alla kanske inte kommer eller skall analysera allt. Det som användarna beskriver att se utifrån sina egna behov och börja där kan vara en lämplig väg att gå.

Att alla på sikt kan komma att få tillgång till samma verktyg och möjligheter för analys behöver inte betyda att fördelarna på sikt kommer jämna ut sig helt. Det ställer stora krav på kompetens och uppfinningsriktighet att hitta vad man faktiskt skall analysera. Det är viktigt att se till sitt eget behov och bevaka området utifrån sina egna förutsättningar för att utnyttja möjligheten att se tidigare okända faktorer vilka faktiskt påverkar och driver måluppfyllnad. Aktörer som är snabba på att anamma en teknologi kan använda det som strategisk konkurrensfördel så länge den stora massan och konkurrenter inte lyckas följa dem. Att bibehålla dessa kommer dock att bli allt svårare allt eftersom färdiga paketeringar och software as a service tjänster sprider sig.

Det är även viktigt att tänka på att analysen aldrig blir bättre än datan som loggas. Det spelar ingen roll vilken analys man gör om man inte har riktigt underlag och det är tänkbart att risken för "dålig" information blir större ju fler källor man inkluderar. Det är därför viktigt att kvalitetssäkra den information som faktiskt används som viktigt underlag.

6.4 Förslag fortsatt forskning

I och med att företeelsen Big Data är såpass nytt skulle det vara mycket intressant att göra en liknande studie om några år för att se om, eller hur, den breda massa tagit till sig tekniken och dess möjligheter. Vilka företag och branscher använder teknologin? Har teknologin växt in i det större området BI eller är de fortfarande separata?

Utföra djupare fallstudier i organisationer som använder sig av den här typen av lösningar för att studera hur de arbetar praktiskt och hur deras ekonomistyrning påverkas av möjligheterna till nya typer av analyser.

Göra en studie i syfte att beskriva hur analysarbetet i organisationer utvecklats med hjälp av IT-stöd i ett större perspektiv, till exempel i relation till Skirletz modell över operativt och strategiskt arbete med BI.

Undersöka hur företag organiserar ägandet av den information som samlas in och problematik som kan uppstå då det finns flera intressenter.

7. Referenslista

Abrahamson, E. (1996). Management fashion. *The Academy of Management Review*, Vol. 21 ss. 254-285.

Abrahamson, E. (1991). Managerial fads and fashions: the diffusion and rejection of innovations. *Academy of Management Review*, Vol. 16 n. 3 ss. 586-612.

AFA Försäkring. (2011). *AFA Försäkring i korthet*. Hämtat från <http://www.afaforsakring.se/Om-AFA-Forsakring/AFA-Forsakring-i-korthet/> den 8 Mars 2012

Apache Software Foundation. *Apache Hadoop*. Hämtat från Apache wiki: <http://wiki.apache.org/hadoop/> den 10 03 2012

Apache Software Foundation. *What is Apache Hadoop?* Hämtat från [hadoop.apache.org: http://hadoop.apache.org/#What+Is+Apache+Hadoop%3F](http://hadoop.apache.org/#What+Is+Apache+Hadoop%3F) den 05 03 2012

Ax, C., & Bjørnenak, T. (2011). Towards a Dynamic Perspective on Management Accounting Innovation. *Working paper, School of Business and Economics and Law at the University of Gothenburg and Norwegian school of Economics and Business Administration, Bergen*.

Ax, C., Johansson, C., & Kullvén, H. (2009). *Den nya ekonomistyrningen*. Malmö: Liber.

Baars, H., & Kemper, H. (2008). Management Support with Structured and Unstructured Data- An Integrated Business Intelligence Framework. *Information Systems Management*, Vol. 25 ss. 132-148.

Birkinshaw, J., Hamel, G., & Mol, M. (2008). Management innovation. *Academy of Management Review*, Vol. 33 no. 4 ss. 825-845.

Bitterer, A. (2011). *Hype Cycle for Business Intelligence 2011*. Gartner Research ID:G00216086.

Blechar, M., Adrian, M., Friedman, T., Roy Schulte, W., & Laney, D. (2011). *Predicts 2012: Information Infrastructure and Big Data*. Gartner Research ID:G00226066.

Broxte, C. (Intervju den 6 Mars 2012). SAS Institute.

Carr, N. G. (2003). IT Doesn't Matter. *Harvard Business Review*, May ss. 41-49.

Casonato, R., White, A., Beyer, M., Adrian, M., Friedman, T., Blechar, M. (2011). *Top 10 Technology Trends Impacting Information Infrastructure, 2012*. Gartner Research ID number:G00227024.

Dalen, M. (2008). *Intervju som metod*. Malmö: Gleerups utbildning.

- Dijcks, J.-P. (2011). *Big Data for the Enterprise*. Redwood shores: Oracle.
- Eriksson, M.-O. (Intervju den 12 Mars 2012). King.
- Fenn, J., & Raskino, M. (2011). *Understanding Gartner's Hype Cycles*. Gartner Research ID: G00214001.
- Forrester Research Inc. *About Forrester*. Hämtat från [forrester.com](http://www.forrester.com/home#/aboutus): <http://www.forrester.com/home#/aboutus> den 20 02 2012
- Gartner Inc. *About us*. Hämtat från [gartner.com](http://www.gartner.com/technology/about.jsp): <http://www.gartner.com/technology/about.jsp> den 17 02 2012
- Gartner, I. (2012). *IT Glossary*. Hämtat från Gartner IT Glossary: <http://www.gartner.com/it-glossary/business-intelligence-bi/> den 10 03 2012
- Google. *Google Trender*. Hämtat från Webbsökning Intresse:Big Data: <http://www.google.com/insights/search/#q=%22big%20data%22&date=4%2F2010%2024m&cmpt=q> den 16 03 2012
- Herschel, G. (2010). *There are few synergies between sentiment analysis and predictive analysis*. Gartner Research ID: G00206413.
- Holme, I., & Solvang, B. (1997). *Forskningsmetodik: Om kvalitativa och kvantitativa metoder*. Lund: Studentlitteratur.
- IBM. (2012). *Bringing big data to the enterprise*. Hämtat från [ibm.com/bigdata](http://www-01.ibm.com/software/data/bigdata/): <http://www-01.ibm.com/software/data/bigdata/> den 05 03 2012
- IBM media relations. (2011). *Vestas Wind Systems turns to IBM Big Data Analytics for Smarter Wind Energy*. Hämtat från <http://www-03.ibm.com/press/us/en/pressrelease/35737.wss> den 14 Mars 2012
- Informatica. (den 27 Januari 2011). *Station Casinos*. Hämtat från <http://www.informatica.com/us/customers/station-casinos.aspx> den 8 Mars 2012
- Jacobsen, D. (2002). *Vad, hur och varför: om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*. Lund: Studenlitteratur.
- Jerresand, J. (Intervju den 9 Mars 2012). Affecto.
- Johnson, R. (den 21 Juli 2010). *Scaling Facebook to 500 Million Users and Beyond*. Hämtat från [facebook.com](http://www.facebook.com/note.php?note_id=409881258919): http://www.facebook.com/note.php?note_id=409881258919 den 10 Mars 2012
- Kaplan, R. (1983). Measuring manufacturing performance: A new challenge for managerial accounting research. *The Accounting Review*. no. 4 s. 686-705

- King, R. (2011). Getting a Handle on Big Data with Hadoop. *Bloomberg Businessweek* 7 september 2011.
- Kjellberg, F. (Intervju den 01 Mars 2012). Acando.
- Kobielus, J. (2012). *The Forrester Wave: Enterprise Hadoop Solutions 2012*. Orlando: Forrester Research.
- Komorowski, M. (den 24 Juli 2009). *A History of Storage Cost* . Hämtat från <http://www.mkomo.com/cost-per-gigabyte> den 10 Mars 2012
- Lämmel, R. (Oktober 2007). Google's MapReduce Programming Model- Revisited. *Journal Science of Computer Programming Vol. 68 u. 3* .
- Laney, D. (2001). *3D Data Management Controlling Data Volume Velocity and Variety*. Stamford: META Group Inc.
- Lidberg, S. (Intervju den 29 Februari 2012). Microsoft.
- Lincoln, Y., & Guba, E. (1989). *Fourth generation evaluation*. Newbury park: Sage publications.
- Loglogic. (2012). *Big Data and cloud survey*. San Jose: Loglogic Inc.
- Malmi, T., & Brown, D. (2008). Management control system as a package- Oppurtunities challenges and research directions. *Management Accounting Research*, Vol. 19 ss. 287-300.
- McKinsey Global Institute. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Seoul: McKinsey & company.
- Merchant, K., & W, V. d. (2007). *Management Control Systems*. New York: Financial Times/Prentice Hall.
- Microsoft News Center (2012). *Microsoft expands data platform*. Hämtat från <http://www.microsoft.com/Presspass/press/2011/oct11/10-12PASS1PR.msp> den 7 Mars 2012
- Microsoft News Center (2012). *Facts about Microsoft*. Hämtat från http://www.microsoft.com/presspass/inside_ms.msp den 2 Mars 2012
- Neubauer, P. (den 12 Maj 2010). *Graph Databases, NOSQL and Neo4j*. Hämtat från InfoQ: http://www.info.com%2Farticles%2Fgraph-nosql-neo4j&sa=D&sntz=1&usg=AFQjCNEAN2FibIQluD-_B3D7LkmvtoyeRA den 05 03 2012
- Pant, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. (ss. 79-86). Philadelphia: Assosiation of computational linguistics EMNLP.

- Patel, R., & Davidsson, B. (2003). *Forskningsmetodikens grunder: att planera, genomföra och rapportera en undersökning*. Lund: Studentlitteratur.
- Pollock, N., & Williams, R. (2010). The Business of Expectations: How Promissory Organisations Shape Technology & Innovation. *Social Studies of Science*, Vol. 40 august ss. 525-548.
- Porter, M. (1996). What is strategy? *Harvard Business Review* , Nov/Dec ss. 61-78.
- Porter, M., & Millar, V. (den Vol. 63 Issue 4 Juli/Augusti 1985). How information gives you a competitive advantage. *Harvard Business Review* , s. 149.
- Prentice, S. (2011). *CEO Advisory: "Big Data" Equals Big Opportunity*. Gartner Research ID: G00211628.
- Rhodin, M. D. (Seminarie den 06 Mars 2012). *CeBit 2012: Building smarter business in the Era of Big Data*. Hannover.
- Rogers, S. (den 1 September 2011). *Big Data is Scaling BI and Analytics*. Hämtat från information-management: http://www.information-management.com/issues/21_5/big-data-is-scaling-bi-and-analytics-10021093-1.html den 08 03 2012
- Russom. (2007). *BI Search and Text Analytic*. Chatsworth: The Data Warehousing Institute.
- Russom, P. (2011). *TDWI Best practices report: Big Data analytics q4 2011*. The Data Warehousing Institute.
- Sallam, R., Richardson, J., Hagerty, J., & Hostmann, B. (2011). *Magic Quadrant for Business Intelligence Platforms*. Gartner Research ID: G02110036.
- Samuelson, L. A., & Olve, N. G. (2001). *Controllerhandboken*. Malmö: Liber.
- SAS Institute AB. (December 2008). *AFA Försäkring förebygger skador*. Hämtat från <http://www.sas.com/offices/europe/sweden/pdf/afa-forsakring-forebygger-skador.pdf> den 8 Mars 2012
- SAS Institute. (u.d.). *Fakta om SAS Institute*. Hämtat från <http://www.sas.com/offices/europe/sweden/omsas/faktaomsas.html> den 3 Mars 2012
- SAS Institute. (2009). *Text Mining for Safety*. Cary: SAS Institute.
- Shilakes, C., & Tylman, J. (1998). *Enterprise Information Portals*. New York: Merrill Lynch.
- Simons, R., Davila, A., & Kaplan, R. (2000). *Performance measurement & control systems for implementing strategy*. Upper Saddle River: Prentice Hall.
- Skrlertz, R. (2002). *Operations-Focused BI vs. Strategic BI*. New York: Information-management.

The Data Warehouseing Institute. *What we do*. Hämtat från tdwi.org:
<http://tdwi.org/pages/about-tdwi/tdwi-business-intelligence-and-data-warehousing-education-and-research.aspx> den 20 02 2012

The Economist. (2010). All too much- Monstrous amounts of data. *The Economist* , 27 februari ss 6.

Tredinnick, L. (2006). Web 2.0 and Business. *Business Information Review* , 23 (4), 228-234.

Twitter. (den 30 Juni 2011). *blogg.twitter*. Hämtat från twitter:
<http://blog.twitter.com/2011/06/200-million-tweets-per-day.html> den 10 03 2012

Vestas Wind Systems. (2011). *New supercomputer among world's most powerful*. Hämtat från <http://www.vestas.com/en/media/news/news-display.aspx?PID=0&NewsID=2708&action=3> den 14 Mars 2012

Zikopoulos, P., Eaton, C., deRoos, D., Deutsch, T., & Lapis, G. (2012). *Understanding Big Data*. New York: McGraw-Hill.