



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Big Data i den digitala matbutiken

En kvalitativ studie om dataanvändning i svensk dagligvaruhandel

Kandidatuppsats

Industrial and Financial Management

VT 2018

Handelshögskolan vid Göteborgs Universitet

Författare

Sebastian Agerhäll 19950322

Christian Schultz 19900811

Handledare

Ove Krafft

Sammanfattning

Dagligvaruhandeln på internet blomstrar och slår årliga försäljningsrekord vilket inneburit att de svenska handlarna ökat sin närvaro på internet. Den ökade aktiviteten på webben medför ökad tillgång på stora datamängder, så kallad Big Data. Detta arbete syftar till att beskriva hur svenska dagligvaruhandlare använder Big Data idag samt vilka möjligheter och utmaningar det finns inför framtiden. Studien har genomförts med en kvalitativ metod där aktörer inom svensk dagligvaruhandeln har studerats genom primärdata i form av intervjuer och sekundärdata i form av podcasts och artiklar. Insamlad data har sedan jämförts med teorier kring dataanvändning samt med en aktör på den brittiska dagligvarumarknaden. Studien konstaterar att de svenska dagligvaruhandlarna är i ett tidigt skede angående användandet av Big Data jämfört med den brittiska, men även att det är ett område studerade aktörer vill utveckla. Studien indikerar vidare att användningen av Big Data för den svenska marknaden troligen kommer att röra sig mot ett ökat fokus på optimering och individanpassning. Studien belyser bland annat Big Data nudges, automatisering och prediktion av efterfrågan som möjligheter för Big Data i dagligvaruhandeln, och huvudsakliga utmaningar identifieras som lågt engagemang från ledningen samt brist på kompetens. Inför framtiden tror vi att företagen bör satsa på att utveckla sitt användande av Big Data men att det bör ske med en tydlig strategi och ett kritiskt tänkande.

Nyckelord: Dagligvaruhandel, Big Data, Artificiell intelligens, e-handel, maskininlärning

Abstract

The online market for fast moving consumer goods (FMCG) is booming and Swedish FMCG-retailers are now increasing their presence online. As a result, the retailers now have a greater than ever access to large quantities of data, so called Big Data. This study aims to describe how the Swedish FMCG-retailers currently manage and use Big Data, as well as identify opportunities and challenges for the future. The study has been conducted using a qualitative method where primary data for Swedish FMCG-retailers has been gathered from an interview, and secondary data has been gathered from podcasts and articles. This data has then been compared with theories about use of Big Data as well as with an FMCG-retailer on the British market. The study concludes that the Swedish retailers are in an early stage of Big Data utilization compared to the British retailer, but also that it's an area the Swedish retailers would like to develop. Additionally, the study indicates that the utilization of Big Data on the Swedish FMCG-market is likely to move towards individualization and process optimization. Big Data nudges, automation and demand prediction are highlighted as opportunities for Big Data, while lack of interest from management as well as lack of required competence are highlighted as the main challenges.

The authors strongly believe that the companies should invest in developing their use of Big Data, but that it should be done with a clear strategy and critical thinking.

Key words: fmcg, Big Data, Artificial Intelligence, e-commerce, machine learning

Innehållsförteckning

1	Introduktion	1
1.1	Bakgrund	1
1.2	Problemdiskussion.....	2
1.3	Syfte	4
1.4	Forskningsfrågor	4
1.5	Avgränsningar.....	4
2	Teoretisk referensram	5
2.1	Big Data	5
2.2	Användning av Big Data	7
2.3	Utmaningar med Big Data.....	8
2.4	Utvärdering av dataarbete.....	10
2.5	Teorisammanfattning	12
3	Metod	13
3.1	Metodval	13
3.2	Datainsamling.....	14
3.2.1	Intervjudesign.....	15
3.2.2	Urval	15
3.2.3	Intervjugenomförande.....	16
3.2.4	Sammanställning.....	17
3.3	Referensram och diskussion	17
3.4	Validitet och reliabilitet	18
3.5	Metodreflektion	18
4	Empiri.....	20
4.1	Företagspresentation	20
4.2	Datainsamling.....	21
4.3	Dataanvändning	22
4.4	Framtidssyn	24
4.5	Utmaningar	26
5	Analys.....	29
5.1	Datakvalitet.....	29
5.2	Dataanvändning	30
5.3	Mognad	31

5.4 Vägen framåt.....	33
5.4.1 Syfte och mål med dataanvändningen	33
5.4.2 Kompetens.....	34
5.4.3 GDPR	35
5.4.4 Logistik.....	35
6 Slutsatser	37
7 Vidare forskning	39
Referenslista	40
Bilaga 1: Frågemall	i

Ordlista

Big Data: Syftar på digitalt lagrad information som skapas i så pass hög hastighet, variation och volym att den är svår att bearbeta med traditionella databasmetoder som hårddiskar (McAfee & Brynjolfsson 2012)

Artificiell Intelligens (AI): Syftar på ett datorprogram som försöker efterlikna förmågor som tidigare setts som unika för mänsklig intelligens. Begreppet innefattar bl.a. maskininlärning och botar där digitala assistenter som Apples Siri är ett exempel på AI. (Flasiński, 2016)

Maskininlärning (ML): Underkategori till AI som handlar om att träna datorer att göra väldigt specifika uppgifter. Bild- och röstigenkänning är exempel på maskininlärningsapplikationer. (Flasiński, 2016)

Data Mining/Analytics: Handlar om att hitta mönster i data eller att utvinna information ur data. Ofta krävs stora datamängder, d.v.s. Big Data, för att den extraherade informationen ska vara pålitlig. (Dahr, 2013)

Data scientist: Vanlig titel på de som arbetar med att utveckla modeller och algoritmer för maskininlärning och Data Mining. (Dahr, 2013)

Internet of Things(IoT): Syftar till vardagsföremål som är internetuppkopplade. I princip allt kan omfattas, exempelvis smarta kylskåp som berättar vad som behöver inhandlas eller smart belysning som kan tändas och släckas på obegränsade avstånd via en mobiltelefon. (ITU, 2012)

GDPR: En personuppgiftslag som träder i kraft i maj 2018 som omfattar alla företag som har verksamhet eller kunder i Europa. Lagen ställer höga krav på transparens i hur data används samt begränsar möjligheterna att lagra persondata. (Datainspektionen, n.d)

BI: Business Intelligence är analys av data som rör affärsaspekter. Maskininlärning och Analytics kan användas för att ta fram BI. (Schmarzo, 2016)

1 Introduktion

1.1 Bakgrund

Digitaliseringen av samhället går snabbt framåt och i takt med att e-handeln kontinuerligt slår ny rekordförsäljning¹ tycks det finnas både incitament och press på företag att öka sin närvaro på internet. För många företag innebär det en omställning som sällan är enkel, och DIBS(2017) beskriver ytterligare utmaningar i hur dagens konsument är målmedveten och utan tålamod när det kommer till osmidiga köpupplevelser. Marknaden har dock blivit allt bättre på att tillgodose kundernas behov(ibid), vilket reflekteras inte minst i mathandeln online som växte med 30% under 2016 (Svenska Digital Handel, 2016). Tillväxtprognosen för 2017 är även den 30% vilket motsvarar en total omsättning om 7,2 miljarder, men trots stark tillväxt motsvarar det endast ca 2,4% av total livsmedelsförsäljning(ibid). Den främsta anledningen till att svenskar handlar mat på internet är för att det är bekvämt och många leverantörer arbetar därför med precisa och flexibla leveranser (DIBS, 2017). Detta har bidragit till en klart uppåtgående trend av matköp på internet(ibid), och för den som ytterligare kan förbättra upplevelsen finns det till synes stor potential.

Bernard Marr (2016) beskriver hur i princip allt vi gör online lämnar spår som kan samlas, aggregeras och analyseras för att nå nya insikter vilket påverkar hur vi handlar med varandra och ge upphov till nya möjligheter. Möjligheterna kan ibland tyckas angränsa till science-fiction med visioner om att varor levereras innan de ens är påtänkta, men faktum är att verkligheten inte är särskilt långt efter. Amazon patenterade till exempel en algoritm för att förutspå handel och fraktdestination baserat på tidigare kundaktivitet redan 2013(Amazon, 2013), och allt fler företag försöker hitta nya sätt att ta vara på de enorma mängder med information som finns om deras kunder, ofta samlat under namnet Big Data, för att stärka sin position både på och utanför internet.

Big Data är ett samlingsbegrepp för datamängder som är tillräckligt stora och komplexa att det är svårt att hantera dem med traditionella databasmetoder, mycket på grund utav mängden

¹ Enligt DIBS rapport *Svensk E-Handel 2017* nådde e-handeln i Sverige ny rekordförsäljning om 109,5 miljarder 2017, vilket är en ökning på nästan 10% från 2016.

ostrukturerad data (Clarke, 2015). Traditionellt har Big Data handlat mycket om att hämta, lagra och analysera stora datamängder, men har på senare tid fått ett tydligare fokus på prediktiv analys, beteendeanalys och mer avancerade metoder som t ex maskininlärning för att utvinna värde ur data (Schmarzo, 2016).

Analys av Big Data har länge varit kopplat till höga kostnader i form av hård- och mjukvara, strukturering av rådata och anställda dataspecialister för att tolka resultaten. Detta innebär att främst kapitalstarka företag kunnat ta del av tekniken, men i takt med att marknaden mognat har allt fler företag som erbjuder ”Big Data as a service” etablerat sig, vilket möjliggör att även mindre företag kan ta del av avancerade algoritmer och analyser. Klarna erbjuder till exempel avancerade algoritmer för att förhindra bedrägerier i tillägg till en betalningslösning (Marr, 2016), och såväl Facebook som Google erbjuder möjligheten att ta del av riktade marknadsföringsstrategier baserade på Big Data. I livsmedelsbranschen är det ofta Tesco som lyfts som exempel för vilka nyttor som användning av Big Data kan medföra. Tesco har bland annat använt Big Data för att optimera sin värdekedja vilket resulterat i minskat slöseri om 30 miljoner pund, reducerat lagerkostnader med 50 miljoner pund och 30% färre varubrister (Miller, 2012). Sammantaget finns en stor mängd möjligheter med Big Data, men det är inte alltid tydligt vilka som tillför något ekonomiskt värde.

1.2 Problemdiskussion

Digitaliseringen för med sig stora möjligheter, men också stora krav på anpassning och förändring hos många företag. Alla förändringsprojekt är riskabla eftersom något känt lämnas för något okänt, men det kan dock vara ännu mer riskabelt att vänta för länge att följa med trenderna. Arbetet med att anpassa verksamheten till en digital värld är därför mycket viktig för de allra flesta företag. Detta återspeglas bland annat i en undersökning av McKinsey & Co (Bughin et al. 2018) där endast 8% av de tillfrågade företagen trodde att deras nuvarande affärsmodeller skulle vara lönsamma om branschen fortsätter utvecklas med samma hastighet. Behovet av förändring i allmänhet, och digitalisering i synnerhet, tycks därmed vara stort.

Det går inte att styra det som inte går att mäta (McAfee & Brynjolfsson, 2012) vilket gör Big Data till en viktig del av digitaliseringen genom enorma informationsmängder. McAfee och Brynjolfsson beskriver hur Big Data har potential att transformera hela företag genom nya sätt att mäta och styra, och hur dessa nya insikter kan bidra till nya konkurrensfördelar. Nyttan ska

dock inte överdrivas, utan det snarare handlar om konkurrensnödvändigheter än konkurrensfördelar när det kommer till informationssystem (Stockdale & Standing, 2005). Big Data kräver dessutom ofta en hel del omställningar, t.ex. anställning av data scientists och ny infrastruktur för insamlande av data. Dessa omställningar är långt ifrån enkla vilket reflekteras av att en stor mängd IT-projekt misslyckas, och många har svårt att få ut det värde som förväntats (Gulla, 2012). Big Data för alltså med sig flera olika utmaningar, varav några av de största är implementering och att hitta rätt data och rätt insikter för att skapa värde för verksamheten.

Dagligvaruhandeln är en väletablerad industri som har funnits mycket länge och består av ett fåtal stora aktörer i Sverige. Affärsmodeller med stormarknader och franchisebutiker har fungerat bra i många år, men under de senaste åren har det dykt upp konkurrenter som endast opererar online vilket förändrat förutsättningarna på marknaden. Mat online är i sig ingen nyhet, men de stora satsningarna har kommit relativt sent jämfört med andra branscher. Andreas Svensson på handelns utredningsinstitut (citerad i SvD 2017) tror att det beror på de stora kedjornas misslyckade satsningar på e-handel i början av 2000-talet vilket medförde att intresset var lågt när tekniken mognat och boomen kom på riktigt. Svensson beskriver vidare att de sena satsningarna kan bero på en ovilja till förändring när den traditionella handeln är så lönsam, men att de samtidigt inte vill hamna på efterkälken att möta morgondagens konsument. Marknaden online är dock fortfarande relativt liten i en bransch där lönsamhet bygger på volym, och de flesta e-handelsbolag inom dagligvaruhandeln gick med förlust 2016 enligt SvDs (2017) granskning.

Förluster till trots så växer marknaden online och framtidstron är stark, men för att lyckas krävs såväl försäljningsvolym som förmåga att hantera helt nya frågor. I en redogörelse över hur Big Data används hos Tesco, Storbritanniens största dagligvarukedja, identifierar Bernard Marr (2016) de huvudsakliga utmaningarna som förändrat kundbeteende, logistik, minskat slöseri med resurser och att möta konkurrensen från framväxande företag. Lösningar på samtliga utmaningar ligger enligt Marr (2016) i realtidsanalys och den allt växande mängden tillgänglig kunddata tack vare Internet of Things. Enligt DalleMule och Davenport (2017) har företag inte kommit långt i användandet av data och mindre än hälften av strukturerad data används för att fatta beslut. Vidare används mindre än 1% av ostrukturerad² data alls (ibid)

² Data i form av exempelvis bilder, ljud, dokument och video som inte är del av en databas

vilket väcker frågor om Big Data är mycket mer än ett modeord i nuläget. Möjligheterna finns onekligen där oavsett om verksamheten sker helt eller delvis på internet, men ännu råder viss ovisshet om de faktiskt tas till vara på, om det skiljer sig mellan verksamheter i samma bransch och i sådana fall vad eventuella skillnader beror på.

Dagligvaruhandeln blomstrar i flera länder och många olika marknader kan vara intressanta, men med tanke på studiens begränsade omfattning och resurser kommer vi huvudsakligen att fokusera på den svenska dagligvaruhandeln. Eftersom den svenska dagligvaruhandeln domineras av ett fåtal stora aktörer medför denna avgränsning att en relativt bred bild kan uppnås med begränsade resurser. Studien kommer ha ett övergripande strategiskt perspektiv på hur Big Data kan användas snarare än ett tekniskt. Detta då ett tekniskt perspektiv snabbt blir mycket komplicerat och för omfattande att undersöka med de resurser vi har tillgängliga. Vidare beskrivs framförallt användning av kund- och persondata, då denna typ av data innehåller mycket potential och GDPR medfört förändrade krav på hur data ska hanteras.

1.3 Syfte

Syftet med studien är att beskriva hur Big Data används inom svensk dagligvaruhandel och ge en överblick av hur användningen kan se ut i framtiden.

1.4 Forskningsfrågor

Samtliga nedanstående frågor ställs med avseende på svensk dagligvaruhandel

- Hur används Big Data i nuläget?
- Vilka risker och möjligheter finns med Big Data?
- Hur kan användningen av Big Data se ut i framtiden?

1.5 Avgränsningar

- Svensk dagligvaruhandel exklusive måltids- och speciallösningar
- Strategiskt perspektiv från aktörernas synvinkel
- Fokus på användning av kund- och persondata

2 Teoretisk referensram

2.1 Big Data

Med Big Data avses generellt stora digitala datamängder som kan genereras från allt från inlägg på sociala medier, till väderleksrapporter till sensordata i fabriker. Definitionen av Big Data är omstridd men innebär ofta att den är för stor för att kunna hanteras av traditionella datahanteringstekniker, t.ex. lagra data på hårddiskar (Clarke, 2015). Även om Big Data är något svårhanterad är det en avgörande förutsättning för avancerade analystekniker och Artificiell Intelligens (AI). Artificiell Intelligens syftar på ett datorprogram som försöker efterlikna förmågor som tidigare setts som unika för mänsklig intelligens (Flasiński, 2016). Likt att människor kan träna upp olika färdigheter och kunskaper måste även de artificiella intelligenserna tränas upp. Det finns många olika typer av AI, men alla har en gemensam grundidé, att försöka anpassa en matematisk modell till stora mängder data, dvs Big Data. Att efterlikna mänsklig intelligens är mycket svårt för matematiska modeller och därför behöver de tränas utförligt vilket innebär att träningsdata måste vara omfattande, Big (Dahr, 2013).

Som ovan nämnt finns det många underkategorier till AI men det finns ingen knivskarp gräns mellan vad som är vad. Några vanliga underkategorier är maskininlärning, Botar och data mining/advanced analytics. Maskininlärning beskrivs ofta som en AI som tränats för att utföra en specifik standardiserad uppgift, exempelvis känna igen bilder på hundar (Flasiński, 2016). Botar, är ofta designade för att interagera med människor, exempelvis chattbotar i kundservice (Machiraju & Modi, 2018). Data mining eller Advanced Analytics som det ibland kallas handlar om att hitta mönster eller korrelationer i data. Detta med syfte att utvinna insikter om data eller att förutspå något (Dahr, 2013).

Data brukar delas in i två kategorier, strukturerad och ostrukturerad data. Strukturerad data är data som från början är sorterad i rader och kolumner med tydliga annoteringar. Strukturerad data är lätt att söka i och att analysera. Det är dock endast en liten del av all tillgänglig värdefull data som är strukturerad. Enligt Seth Grimes (2008) är 80 % av affärsrelevant data ostrukturerad. Ostrukturerad data är all data som inte har någon tydlig inre struktur, exempelvis bilder, textdokument och email. Ostrukturerad data innehåller mycket mer information än strukturerad data men är betydligt svårare att analysera ostrukturerad data.

Både strukturerad och ostrukturerad data kan ha hög kvalitet, och det är snarare informationstätheten och lätthanterligheten som skiljer dem åt. (Coblin et al. n.d)

En mycket viktig del inom dataanalys är att data håller hög kvalitet. Om inte insamlad data representerar verkligheten på ett bra sätt kan inga värdeskapande insikter fås (Clarke 2015). Att samla in all data som går är inte alltid det bästa då det kan innebära att all överflödiga data utan information döljer värdefull data. Dessutom är det dyrt att samla in och lagra data som inte behövs (Marr, 2015c). Clarke (2015) beskriver sju viktiga faktorer för god datakvalitet.

- D1: Syntactical Validity: Hur väl innehållet i data överensstämmer med det verkliga området, på vilket det studerade objektet agerar.
- D2: Appropriate Identity Association: Hög säkerhet på att dataobjektet hör ihop med det verkliga objekt som avses studeras.
- D3: Appropriate Attribute Association: Entydighet kring vilka egenskaper hos ett verkligt objekt som data representerar.
- D4: Appropriate Attribute Signification: Entydighet i vilket tillstånd hos det verkliga objektet som data representerar.
- D5: Accuracy: Noggrannhet i data, ofta representerat av ett konfidensintervall, tex. +/- 5 minuter.
- D6: Precision: Hur detaljerad information om det verkliga objektet som fås av data.
- D7: Temporal Applicability: Klarhet kring när och var data samlades in. Viktigt för data som är starkt beroende av tid och plats. Exempelvis säsongsvarierande data eller demografisk data.

Data behöver inte uppfylla alla kriterier för att vara användbar men genom att utvärdera tillgänglig data utifrån dessa kriterier bildas en uppfattning om vilka insikter data kan ge, hur pålitliga insikterna är och på vilka områden försiktighet kring slutsatser ska iakttas. Om data uppfyller kriterierna D1-D4 är data starkt kopplat till det som avses undersökas och kan då ge mycket information om objektet. Om en eller flera av dessa kriterier inte uppfylls kan det vara svårt att dra slutsatser om kausalitet även om korrelationer finns. D5- och D6- kriterierna behandlar hur exakt information data kan ge. Vad som är tillräcklig exakthet beror på vad syftet med dataanalysen är; om syftet är att förutspå något med särskild precision bör även data vara minst lika detaljerad. (Clarke, 2015)

2.2 Användning av Big Data

Litteraturen nämner flera olika områden där Big Data kan komma till användning inom handeln. Genom att analysera Big Data kan företag till exempel utvinna information och insikter som gör det möjligt att öka kundförståelsen och skapa mer effektiv marknadsföring med riktade erbjudanden (Bose, 2009; Hitchcock, 2017; Marr, 2015b). Såväl Hitchcock (2017) som Marr (2015b) nämner möjligheterna att identifiera och följa värdefulla kunder där analys av köpmönster medför att efterfrågesvängningar kan predikteras. Det är även möjligt att få insikter om när vissa specifika kunder är mer eller mindre benägna att handla vilket marknadsföringen kan anpassas efter.

Carolan (2018) beskriver ytterligare en användning av Big Data i handeln i form av så kallade Big Data nudges. Begreppet nudges beskrivs av Thaler (2008) som subtila positiva förstärkningar som ska "knuffa" individer att ta vissa beslut, exempelvis skriva ut CO2 fotavtrycket på menyer för att få restaurangkunder att välja det mest miljövänliga alternativet. Nudging har länge använts inom handeln, exempelvis genom bakgrundsmusik, olika ljussättningar samt placering av impulsköpsvaror vid kassan. Big Data nudging skiljer sig från klassisk butiksnudging genom att det möjliggör individuell nudging. Det går även att anpassa nudges efter externa förhållanden, exempelvis väder, tidpunkt på dygnet och sociala trender. Genom att korrelera försäljningsdata med vilka nudges som konsumenten mottagit kan systemen lära sig vilken typ av nudges som är effektiva för den specifika individen och när individen är mest mottaglig för dem. Optimerade nudges kan utifrån denna data genereras automatiskt. (Carolan, 2018)

Bose (2009) beskriver hur analys av data kan ge företag en fullständig bild av sin verksamhet, vilket i sin tur innebär att företaget kan optimera sitt beslutsfattande. Beslutsfattande är en vanlig och viktig del av användandet av Big Data eftersom data utgör en bra grund och förstärkning för beslut (Bose, 2009; Hitchcock, 2017; Marr, 2015b; McAfee & Brynjolfsson, 2012). I takt med att den tekniska utvecklingen går framåt och att företag blir allt mer digitala så finns ett allt större underlag för dataanalys genom avancerade analysmetoder. Detta innebär att allt eftersom utvecklingen går framåt så kan det bli möjligt att automatisera vissa beslut som tidigare varit beroende av egna bedömningar (Bose, 2009; McAfee & Brynjolfsson 2012). Ytterligare exempel på användning av Big Data är prisoptimering i realtid,

bedrägeriskydd genom maskininlärning (Marr, 2015b) och optimering av leverantörskedjor (Hitchcock, 2017).

2.3 Utmaningar med Big Data

Stora datamängder är mycket värdefulla och säker datalagring är därför mycket viktigt (Marr, 2015c). I ett omtalat fall 2013 blev den amerikanska matvarukedjan Target bestulna på kredituppgifter för 70 miljoner kunder (BBC, 2014). Detta resulterade i direkta kostnader på 202 miljoner dollar i form av kundkompensationer och juridiska kostnader och skandalen ledde dessutom till ett stort tapp i kundförtroende (ibid).

Förutom vikten av datasäkerhet är även hanteringen av data mycket viktig för företagen. General Data Protection Regulation (GDPR) är en lag som berör alla företag som har verksamhet, inom EU. Den nya lagen lägger stor vikt på transparens vid dataanvändning. Företag måste förklara precis vilken data de avser samla in och vad den ska användas till. Vid dataintrång måste alla individer vars personuppgifter berörs meddelas omgående. Innan persondata börjar samlas in måste godkännande från individerna i fråga erhållas. Individerna har dessutom rätt att begära att få sin data raderad, med undantag för data som behöver lagras enligt överliggande lagar, t.ex. Bokföringslagen. (Datainspektionen, 2018)

Förutom de kraftiga juridiska riktlinjer som företagen måste anpassa sig efter finns det även etiska problem med att lagra och analysera personuppgifter (Marr, 2015c). Kostnaderna vid undermålig hantering av personuppgifter kan sträcka sig förbi de direkta juridiska kostnaderna då Data Privacy är en viktig fråga för konsumenterna. Detta blev tydligt i samband med den så kallade Cambridge Analytica skandalen där Facebook blev starkt kritiserade efter att användardata från 50 miljoner konton läckt till analysbolaget Cambridge Analytica (Meyer, 2018).

Att bygga infrastruktur för att kunna samla in och lagra stora datamängder är dyrt (Marr, 2015c). Det krävs i tillägg dyr kompetens i form av exempelvis data scientists för att hantera och analysera data samt för att säkerställa att ovanstående två punkter uppfylls. Därför hävdar Marr (2015a; 2015c) att det är viktigt med en tydlig och genomtänkt strategi kring hur Big

Data ska användas. Risken är annars att företagen lägger stora pengar på omtalade buzzwords utan att få ut något verkligt värde (Marr, 2015c).

Felaktig analys är en mycket stor risk då det kan leda till att en redan påkostad investering leder till felaktiga insikter som istället för att skapa värde till företaget, skadar företaget genom att fokus läggs på fel saker (Marr, 2015c). Företag måste vara vaksamma för att inte övertolka kausalitet i data. Ofta går det att hitta korrelationer och mönster i data som inte nödvändigtvis måste vara kausala. Marr (2015c) ger exemplet att försäljning kan observeras öka i samband med ett stort sportevenemang, men anledningen till försäljningsökningen kanske inte beror på sportevenemanget i sig utan bara på att det är fler människor i staden än vanligt. Ett musikevenemang skulle kanske ha samma effekt. Det är även viktigt att undvika så kallad confirmation bias, dvs att det finns en förutfattad uppfattning om vad data ska berätta, vilket kan leda till att signaler som motsäger detta ignoreras. (Marr, 2015c)

En omställning till användning av Big Data kräver ökad teknisk kompetens som kan implementera de modeller som används för att analysera data, men för att skapa mycket värde krävs även att ledningen gör Big Data till en del av sin strategiska planering. Ariker et al. (2014) skriver om behovet av "data-tolkare" som ett komplement till den rent tekniska kompetensen. "Data-tolkarnas" uppgift är att vara en brygga mellan olika affärsfunktioner. De har kompetens både inom den specifika affärsfunktionen och inom Big data och kan på så sätt hitta optimala användningsområden för Big data inom just detta området. (Ariker et al, 2014)

Marazzi (2017) menar att ett vanligt förekommande problem inom Big Data i affärsvärlden är dålig kunskap från högsta ledningen. I många fall har företag bestämt sig för att satsa på Big Data utan att ha tänkt igenom varför och anställer data scientists med förhoppning att resten löser sig. Ofta har företagsledare orimliga förväntningar vilket leder till att högvärlödade data scientists tvingas lägga mycket tid på att diskutera med och förklara för ledningen istället för att arbeta med själva dataanalysen. Detta resulterar i att de anställda tröttnar och säger upp vilket leder till höga kostnader för företagen. Dessutom krävs det ofta en ordentlig datainfrastruktur som tar hand om insamling, lagring och städning av data för att det ska vara möjligt att analysera. Det är med andra ord mycket viktigt att ha kunskap om Big Data innan det rekryteras dyra data scientists. (Marazzi, 2017)

2.4 Utvärdering av dataarbete

För att kunna utvärdera användandet av Big Data presenteras här två modeller. Modellerna är tänkta att komplettera varandra där den ena modellen fokuserar på att beskriva ett företags mognad inom Big Data och den andra ger riktlinjer för hur arbetet med Big Data bör implementeras.

Big Data Business Model Maturity Index (BDBMMI) – Bill Schmarzo (2015)

Big Data Business Model Maturity Index (BDBMMI) beskriver fem olika faser som ett företag kan vara i när det gäller hur välutvecklat deras användande av Big Data är (Schmarzo, 2015).

Business Monitoring

Användande av business intelligence verktyg för att i huvudsak övervaka hur det går för företaget genom mycket specifik data angående den faktiska prestandan på en specifik del av företaget. Deskriptiv ansats där historiska data summeras och beskrivs.

Business Insights

Större mängder data från ett bredare spektrum inkluderas för att kunna göra prediktiv analys och skapa insikter om exempelvis konsumentbeteenden, efterfrågemönster eller operativa angelägenheter, som kan användas som stöd i beslutsprocesser. I denna fas handlar det om att ta fram goda verktyg inom analytics som kan ta fram värdefulla insikter ur stora datamängder, exempelvis genom nyckeltal.

Business Optimization

Tolkande av de insikter företaget fått från data för att sedan använda dem till att optimera olika delar av verksamheten. Det är även möjligt att kommunicera ut dessa insikter till kunder för att eventuellt påverka deras konsumentmönster.

Business Monetization

Data används för att skapa nya intäktsströmmar, exempelvis genom att sälja data eller insikter, exempelvis genom trendanalys. Insikter kan även integreras i produkter eller tjänster för att skapa “smarta” produkter och tjänster som kan leverera nytt värde till kunderna.

Business Metamorphosis

Företaget använder data, analytics och maskininlärning för att ställa om hela verksamheten, gällande exempelvis kunderbjudande, leverantörer eller marknadsföring. Schmarzo (2015) ger

exemplet om företag som övergår från att sälja produkter till att sälja tjänster eller upplevelser, till exempel att Boeing säljer “Air miles” istället för flygplan.

SMART – Bernard Marr (2015a)

SMART (Marr, 2015a) beskriver hur företag bäst tar till vara på den stora datamängd de har tillgänglig för att maximera insikter och nytta av data. Styrkan i SMART ligger i att kunna navigera oceaner av data för att hitta de små områdena som är av betydelse, och består av nedanstående fem steg.

Start with strategy

Istället för att börja med att utforska vilken data som finns tillgänglig bör starten ligga i att etablera vad organisationen vill uppnå. Strategi bör förtydligas rörande såväl verksamheten i stort för enskilda områden, och organisationen bör i tillägg etablera vilken typ av information som behövs för att nå målet.

Measure metrics & data

Big data är endast användbart om det besvarar frågor som organisationen söker svar på, och för att förstå vilka typer av svar som organisationen kan få tillgång till, måste de också förstå vilken data som finns tillgänglig, t ex strukturerad/ostrukturerad, eller extern/intern. Efter att datatyper etablerats bör fokus ligga på att se vilken som bäst kan användas till den strategi som etablerats, men en bra tumregel är att börja med intern och strukturerad data.

Apply analytics

För att kunna extrahera meningsfulla och användbara insikter måste data analyseras, vilket kan ske till exempel genom text- eller bildanalys. Genom dessa kan olika mönster och trender identifieras, och till exempel öka förståelsen för hur organisationens kunder agerar. Transparens är dock viktigt, och kunder bör informeras om hur deras data används.

Report your results

Analyser och de insikter som kan nås genom dem är av potentiellt stor nytta, men endast om de kommuniceras på ett sätt som mottagaren förstår, och att de når rätt mottagare. Marr (2015a) rekommenderar att resultaten rapporteras kortfattat och gärna visuellt, och i tillägg att data som identifierats som fortsatt användbar konverteras till nyckeltal .

Transform your business

Data måste ses som en färskvara och det är viktigt att använda de insikter organisationen uppnår, annars är det svårt att transformera organisationens medarbetare till att arbeta baserat på evidens. Vidare är dataanalys någonting som bör ses som ett konstant pågående projekt och engagemang för att effektivisera implementering av strategi, och organisation bör alltid ha ögonen öppna för nya affärsmöjligheter.

2.5 Teorisammanfattning

Ovanstående teorier kommer användas för att analysera statusen för de svenska dagligvaruhandlarna i dagsläget samt för att utvärdera möjligheter för framtiden. Mer specifikt kommer BDBMMI användas som en måttstock för hur långt företagen har kommit i sitt arbete med Big Data. För att kunna besvara hur Big Data kan användas i dagligvaruhandeln undersöks möjligheterna, men även förutsättningarna för ett lyckat Big Data användande. Med hjälp av teorierna kring datakvalitet (Clarke, 2015) analyseras kvaliteten på data som insamlas idag av dagligvaruhandlarna. Ariker et al. (2014) beskriver vikten av kompetens som förutsättning för lyckad Big Data implementering. Bose (2009), Marr(2015a; 2015c), Carolan (2018) och Hitchcock (2017) beskriver möjligheter och risker med Big Data både generellt inom affärsvärlden och specifikt inom livsmedelshandeln. Utifrån detta analyseras möjliga framtida vägar för de svenska livsmedelshandlarna. SMART-modellen (Marr, 2015a) används för att avgöra hur företagen ska gå tillväga för att få ut den potential som finns i Big Data.

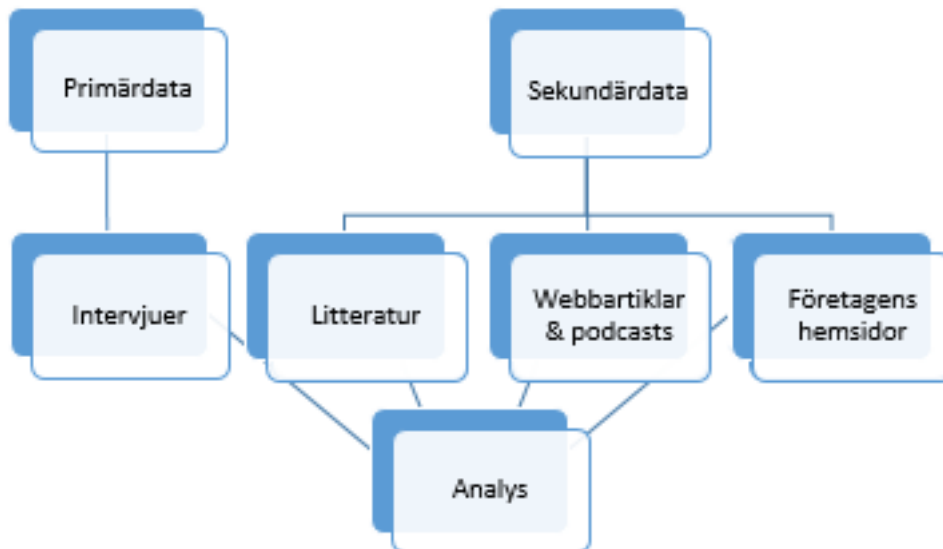
3 Metod

3.1 Metodval

Syftet med studien är att beskriva hur applikationer av Big Data kan användas inom dagligvaruhandeln, med bakgrund av att den ökande e-handeln medför större tillgång till data. Vidare är vår förhoppning att bidra med insikter i hur dataanvändningen ser ut i nuläget, och om hur den svenska dagligvaruhandelns användning av Big Data kan komma att se ut i framtiden genom att jämföra med en brittisk aktör som kommit långt i sitt dataarbete. Vi har valt att använda oss av en kvalitativ undersökningsmetod eftersom en sådan är lämplig när problemet handlar om att beskriva ett fenomen (Patel & Davidson, 2014). Arbetssättet har genomgående varit induktivt med fokus på att utforska, hitta och analysera olika mönster vilket i regel är lämpligt vid kvalitativ forskning (Bryman & Bell, 2011; Patel & Davidson, 2014).

Den kvalitativa metoden och den induktiva ansatsen har två huvudsakliga fördelar för studien. Dels är det, som nämnt ovan, ett lämpligt upplägg när det finns begränsad information om ämnet (Jacobsen, 2002), men det medför också en viss flexibilitet genom att problemformuleringen kan justeras beroende på de upptäckter som görs under datainsamlingens gång (Bryman & Bell, 2011; Jacobsen, 2002). Med tanke på att forskningen som finns inom vårt specifika område är begränsad så passar det studien väl, men det innebär samtidigt vissa konsekvenser. Dels är det svårt att bedöma teorins räckvidd men också att det, som ofta med kvalitativ forskning, saknas en väl etablerad metod (Patel & Davidson, 2014). Det innebär dock inte att vi arbetat förutsättningslöst, utan snarare att det i stort inte kommer gå att generalisera resultaten.

3.2 Datainsamling



Figur 1: Överblick av vår datainsamling

Datainsamlingen skedde i två huvudsakliga steg med start i en litteraturstudie för att skapa en uppfattning om teoretiska synpunkter på hur Big Data kan användas i dagligvaruhandeln. Vidare genomfördes insamling av primär- och sekundärdata med syfte att ge en uppfattning om hur situationen på den svenska marknaden ser ut. Förhoppningen rörande primärdata var att kunna genomföra intervjuer med representanter från de stora aktörerna på den svenska marknaden, eftersom intervjuer ger möjligheten att dra slutsatser baserade på respondenternas erfarenheter och perspektiv (Bryman & Bell, 2011). Det visade sig dock svårt att genomföra intervjuer i någon större utsträckning eftersom flera tillfrågade respondenter valde att avstå, med tidsbrist som anledning. Vi har därför genomfört ett gediget arbete relaterat till sekundärdata där företagens hemsidor och datapolicys har granskats för att ge en blick av hur data samlas och används. Vidare har podcasts där representanter från de olika aktörerna medverkat används eftersom dessa, trots en mer övergripande nivå, bedöms ge en tillräckligt god bild av arbetet med synen på e-handel, arbetet med data och framtidens livsmedelshandel.

3.2.1 Intervjudesign

Vid frågeformuleringen gjordes en avvägning mellan hög och låg grad av strukturering och standardisering varpå valet föll på ett upplägg som Patel och Davidson (2014) beskriver som semistrukturerat. Anledningen är att en låg grad av strukturering och standardisering innebär möjlighet till öppna frågor som i sin tur möjliggör en kvalitativ analys av resultaten (ibid). Detta var viktigt för studien eftersom den syftar till att beskriva faktiska förhållanden utan tidigare antaganden, men viss struktur i form av en frågemall innebar i sin tur en försäkran om att vi faktiskt täckte in områden som ansågs viktiga. Detta upplägg gav i tillägg respondenten frihet att utforma svar på sina egna sätt samtidigt som vi som undersökare kunde ställa följdfrågor för att öka detaljnivån (Bryman & Bell, 2011).

Frågemallen (bilaga 1) utformades som en strukturerad lista med ett antal huvud- och underfrågor att använda som stöd under intervjuer. Huvudfrågorna baserades dels på vårt syfte och frågeställning, men också områden som ansågs relevanta för att stödja analys och diskussion. Underfrågorna är i sin tur sekvenserade enligt tratt-teknik eftersom det både är motiverande och aktiverande för respondenten (Patel & Davidson, 2014) och säkerställer tillräcklig information.

3.2.2 Urval

Primärkällor

Tänkta respondenter på den svenska marknaden valdes genom att vi sökte efter personer med positioner specifikt inom Big Data och maskininlärning eller mer övergripande i form av e-handelschefer. Anledningen till denna form av urval baserades på att vi antog att den information vi sökte inte kunde säkras genom ett slumpmässigt urval, vilket gör ett målstyrt urval likt vårt mer lämpligt (Bryman & Bell, 2011). Respondenten för det brittiska företaget valdes i sin tur genom ett bekvämlighetsurval (ibid) och kontaktades på rekommendation av en gemensam bekant. Vi hade ingen särskild kännedom om företaget vid kontakttillfället, och det var främst under dagarna innan intervjun som det blev tydligt vilken roll det faktiskt har på den brittiska marknaden. Totalt resulterade våra kontakter med företag i endast två intervjuer men med hänsyn till studiens begränsade omfattning, mängden aktörer, och tillgången på sekundärdata ansågs svaren enligt vår bedömning vara tillräckliga.

Sekundärkällor

Eftersom mängden primärdata var begränsad gjordes flera ansträngningar för att säkra material i form av sekundärdata efter intervjuerna var genomförda. Denna bestod i huvudsak av podcasts och artiklar med intervjuer av nyckelpersoner inom e-handel och dataanalys hos Coop, ICA och Axfood. Urvalet av podcasts, se tabell 1, gjordes huvudsakligen baserat på innehållsbeskrivning där vi fokuserade på ämnen rörande e-handel eftersom det i många fall berör frågor som kan kopplas till Big Data, exempelvis individualisering och integritet. Användandet av podcasts medförde att vi kunde nå insikter rörande hur de olika aktörerna arbetar med dessa områden som annars inte vore möjliga, och innebar likt en inspelning även att svaren kan återges exakt. Samtidigt så saknades dock möjlighet att ställa följdfrågor, och eftersom podcasten inte handlade specifikt om Big Data så medförde det att vi inte fick tillgång till det djup vi önskat. Samtliga podcasts är från E-handelstrender och leds av författaren och föreläsaren Urban Lindstedt. Vi upplevde inte Lindstedt som partisk och intervjudeltagarna fick komma väl till tals, men eftersom han har ett stort intresse av e-handel så kan det ha påverkat det kritiska perspektivet, även om utmaningar med e-handel diskuterats.

Person	Organisation	Position	Avsnitt
Anna Åkesson	Hemköp(Axfood)	E-handelschef	44
Claes Hassel	COOP	E-handelschef	44
Magnus Rosén	ICA	E-handelschef	52

Tabell 1: presentation av använda podcasts

3.2.3 Intervjugenomförande

Intervjun med Axfood genomfördes genom att frågemallen skickades via mail eftersom det inte fanns möjlighet att få in en intervju via telefon eller i person med tanke på respondentens pressade tidsschema. Detta har fått viss inverkan på resultaten i form av att respondenten inte velat svara på vissa frågor, vilket förvisso kan vara oberoende av intervjusätt, men också i form av att dialog och öppen diskussion saknas. Samtidigt innebär dock svar i skrift att respondentens svar kan registreras exakt och svaret gav tillräckligt med information för att

vara av värde för studien. Vidare fanns respondenten tillgänglig för ett antal följdfrågor via mail vilket medförde en relativt djup insikt i Axfoods arbete.

Intervjun med den brittiska aktören genomfördes via Skype vilket hade två fördelar. Dels innebar det att vi kunde nå respondenten, som var på resande fot, och dels innebar Skype att intervjun kunde genomföras med både ljud och bild vilket innebar en mer personlig kontakt trots att vi inte var på samma plats rent fysiskt. Inför intervjun skickades de övergripande frågorna över och respondenten informerades om intervjuens roll i studien, rätt till den anonymitet som respondenten begärt samt möjlighet att avbryta samarbetet utan orsak. Vid själva tillfället spelades intervjun in med respondentens godkännande, vilket gjorde att vi kunde fokusera på att vara närvarande istället för anteckningar, vilka fördes endast i begränsad omfattning. Information om studien, hur materialet kommer användas, rätt till anonymitet och så vidare ingår i de forskningsetiska principer som Bryman och Bell (2011) lyfter, och bidrar till att säkerställa att studien håller god forskningsetisk standard.

3.2.4 Sammanställning

Intervjuer och podcasts transkriberades och vi diskuterade därefter respondenternas olika svar och vad som ansågs mest relevant för studiens syfte och frågeställningar. I det transkriberade materialet försökte vi urskilja olika likheter och samband som var ofta förekommande, med syfte att urskilja mönster i svaren. Vi fann ett flertal likheter, men också skillnader, vilka sedan utgjorde en grund för att diskutera våra frågeställningar i jämförelse med det teoretiska underlaget.

3.3 Referensram och diskussion

Den teoretiska referensramen är vald utifrån artiklar som vi främst sökte efter via universitetsbibliotekets supersök, då denna inkluderar en stor mängd databaser vilket ansågs lämpligt då Big Data kan anta flera olika former och perspektiv. De författare som inkluderas i det teoretiska ramverket valdes utifrån att de är refererade av andra och/eller är välkända profiler inom ämnet vilket kan tyda på legitimitet och en bra grund för studien.

Referensramen fungerar dels för att ge en inblick i ämnet, men huvudsakligen som stöd i diskussion och analys av det empiriska materialet.

3.4 Validitet och reliabilitet

Validitet och reliabilitet kopplas ofta till kvantitativa studier där mätning är en central del (Bryman & Bell, 2011), vilket inte är fallet för denna studie. Vi har därför valt oss att använda oss av två alternativa kriterier, trovärdighet och äkthet, som Guba och Lincoln (1994, refererad i Bryman & Bell, 2011) föreslår för kvalitativa studier. Trovärdigheten inkluderar fyra områden; tillförlitlighet, pålitlighet, konfirmering och överförbarhet, vilka tillsammans med äkthet används för att diskutera studien.

Tillförlitlighet innebär att studien följer de ramverk som finns och att respondenterna tar del av resultaten för att säkerställa att vi uppfattat deras svar korrekt, så kallad respondentvalidering (Bryman & Bell, 2011). För att säkerställa detta har vi följt riktlinjer för kvalitativa metoder och intervjuer som beskrivet ovan, och respondenter har tagit del av det material vi sammanställt för att säkerställa korrekt återgivning.

Pålitlighet innebär enligt Bryman och Bell (2011) att vi lämnar en redogörelse för forskningsprocessens olika beståndsdelar. För att säkerställa detta har vi i metoddelen redogjort för vårt tillvägagångssätt samt bifogat vår intervjumall.

Konfirmering innebär att vi som forskare försöker säkerställa att vi har agerat i god tro i avsaknad av fullständig objektivitet (Bryman och Bell, 2011). Vi har så långt det går försökt att inte påverkas av tidigare erfarenheter och preferenser, vilket dels redovisas i våra redovisade tillvägagångssätt, men också genom strävan för ett kritiskt förhållningssätt till både litteratur och respondenter. Vi har i tillägg reflekterat över hur olika metodval kan ha påverkat resultatet.

Överförbarhet innebär att utvärdera i vilken utsträckning som resultaten kan överföras till en liknande kontext (Bryman & Bell, 2011). Denna studie undersöker användning av Big Data på en väldigt specifik marknad och handlar om djup snarare än bredd, vilket innebär att överförbarheten försämras och inga generella slutsatser kan dras. Det är dock värt att notera att applikationer av Big Data ofta kan anpassas, varpå viss överförbarhet ändå finns.

3.5 Metodreflektion

Använd metod har efter justeringar fungerat överlag väl i avseende till studiens syfte och frågeställning men har samtidigt ett antal begränsningar vilka är både generella, men också specifika för just denna studie. Kvalitativa studier är i allmänhet tids- och arbetskrävande och det saknas ofta ett etablerat arbetssätt, vilket innebär att varje kvalitativ studie tenderar att kräva en egen metod. Detta medförde att vi upplevde det som svårt att hitta referenser till hur arbetet bäst skulle genomföras även om det fanns vissa övergripande teman som till exempel fallstudier och intervjuer.

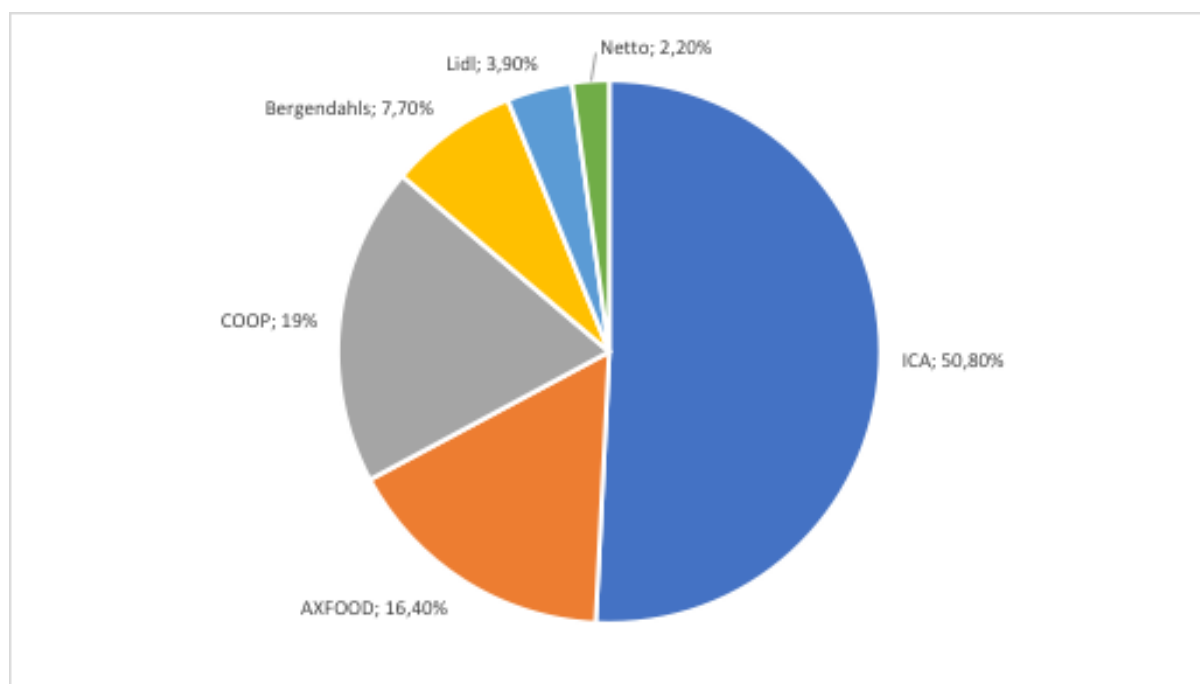
Genom att skicka ut syfte och frågor i förväg kan vi ha påverkat respondenten till att konstruera vissa svar i förväg för att matcha det vi var ute efter, men samtidigt finns det vissa fördelar med att respondenten har en chans att förbereda sig. Dels kan utsvävande svar undvikas, och dels kan respondenten säkerställa att svaren är i linje med företagets riktlinjer vilket minskar risken för att de, trots anonymisering, drar sig ur vid ett senare tillfälle. Anonymisering i sin tur för- och nackdelar i form av att studien eventuellt tappar viss tyngd i avsaknad av information om intervjuobjektet, men samtidigt innebär anonymisering tillgång till information och insikter som vi annars inte haft möjlighet till.

Den största begränsningen i denna studie ligger i att en stor del av resultaten bygger på sekundärdata för den svenska marknaden vilket medför att den djupgående förståelse som intervjun syftade att ge bitvis saknas. Den anledning som nekande respondenter uppgett är tidsbrist, men det kan likväl bero på att respondenterna ej uppfattat det potentiella värdet i deltagandet i studien. Detta beror eventuellt på studiens kvalitativa natur och induktiva metod, men en kvantitativ eller mer teknisk studie ansågs samtidigt för komplext med de resurser som fanns tillgängliga. Trots att mängden primärdata är begränsad tror dock författarna starkt på att den information och de insikter som samlats in och presenterats är värdefulla både ur ett praktiskt och teoretiskt perspektiv - inte minst inför vidare forskning.

4 Empiri

Nedan följer en kort beskrivning av de studerade företagen med information om deras verksamhet i dagsläget. Sedan presenteras vilken data de samlar in, hur insamlingen går till och hur insamlad data används. Slutligen beskrivs hur företagen ser på framtiden, vilka utvecklingsområden de anser viktiga samt vilka utmaningar de står inför.

4.1 Företagspresentation



Figur 2: Marknadsandelar i svensk dagligvaruhandel 2017(DLF, Delfi & HUI Research 2018)

AXFOOD

Axfood är en av största aktörerna inom svensk dagligvaruhandel med en marknadsandel på ca 16 %, se figur 2. Inom Axfood ryms varumärkena Willys, Hemköp, mat.se, Snabbgross, Tempo och Middagsfrid. Några av dessa varumärken har både butik och e-handel, medan mat.se och middagsfrid endast har e-handel. (Axfood, 2018).

COOP

Coop är ett medlemsägt konsumentkooperativ med 655 butiker och en marknadsandel på ca 19 %, se figur 2. Coop har en vision om att vara den goda kraften i svensk dagligvaruhandel och detta försöker de uppfylla genom hårda satsningar på hållbarhet och hälsa (Coop, 2018a).

ICA

ICA är Sveriges ledande aktör inom dagligvaruhandeln (DLF, Delfi & HUI Research, 2018) och samlar ca 1300 butiker som drivs av fria handlare (ICA, 2018a). Detta innebär att kunderna kan mötas av lokalt anpassade koncept, erbjudanden och sortiment samtidigt som skalfördelar kan uppnås genom ett omfattande samarbete (ibid).

Anonym Organisation X

Organisation X, hädanefter ORGX, är en av de stora aktörerna på den brittiska dagligvarumarknaden och erbjuder ett heltäckande sortiment av livsmedel online, vilket kompletteras av varumärken inom såväl smink som hundmat (respondent 1, ORGX). Några fysiska butiker finns inte, utan organisationen använder sig istället av lager där, beroende på lagervariant, upp till 90% av alla aktiviteter är automatiserade. Förutom dagligvaruhandeln har ORGX:s verksamhet utvecklats till att även omfatta ”groceries as a service”(GaaS). GaaS innebär att såväl plattform som automatiserade lager erbjuds och distribueras till andra återförsäljare. (ibid)

4.2 Datainsamling

Undersökningen visar att samtliga företag arbetar med data, att den är av stor vikt i det dagliga arbetet, och att det överlag finns stora likheter mellan både hur den samlas in och av vilka syften oavsett vilken marknad det gäller. Gällande typ av data finns skillnader där respondenten på Axfood beskriver hur de huvudsakligen arbetar med strukturerad data, medan respondenten på ORGX istället beskriver stora investeringar i att samla och lagra både strukturerad och ostrukturerad data, eftersom den någon gång kan komma att bli användbar. Det övergripande syftet är att förbättra kundupplevelsen och de tjänster som erbjuds, och data som används består främst av butikshistorik och persondata, varav den senare är det som studien fokuserat på. Den persondata som samlas in kategoriseras enligt följande:

Kategori	Exempel
Personuppgifter	Kontaktuppgifter, betalningsinformation, ip-adress
Köphistorik	Varor, när och hur de handlats och av vem, inlösta kuponger
Interaktion	Kontakter med kundtjänst, tekniska specifikationer, självscanning, webbnavigering

Tabell 2: kategorier av persondata

Den huvudsakliga datakällan för dessa uppgifter uppkommer av registrering av medlemskap i företagens respektive lojalitetsprogram och/eller hemsida för att kunna utnyttja de tjänster som erbjuds online. Respondenten på ORGX beskriver i tillägg hur datainsamling förekommer i form av datautbyte med personliga assistenter som Amazon Alexa³, och överlag uppger aktörerna fem huvudsakliga syften för datainsamling enligt nedan:

Syfte	Exempel
Köp- och medlemskapshantering	Säkerställa identitet, hantera historik, genomföra betalning och leverans
Kundtjänst	Kommunicera med kunder samt utreda och besvara reklimationsärenden
Utveckling	Göra tjänster mer användarvänliga, utvärdera marknadsföring samt ge underlag för nyetableringar
Säkerhet	Skydda IT-miljön mot angrepp, kontrollera att självscanning används korrekt och fullfölja rättsliga förpliktelser
Anpassning	Direktmarknadsföring, personliga produktrekommendationer

Tabell 3: aktörernas syften med datainsamling

4.3 Dataanvändning

Dataanvändningens övergripande syfte är detsamma som syftet med datainsamling: att förbättra upplevelsen i användandet av de tjänster som respektive aktör erbjuder. Eftersom en upplevelse är någonting subjektivt har utvecklingen gått mot ett ökat fokus på individanpassning vilken drivs av dataanalys. Undersökningen visar på fyra stycken huvudsakliga områden för analys av stora datamängder.

³ Amazon Alexa är en röststyrd personlig assistent som bland kan användas för att exempelvis ställa väckarklocka, spela musik, spela upp nyheter, styra andra apparater i hemmet och så vidare.

4.3.1 Marknadsföring och försäljning

De uppgifter som samlas in vid köp används för att genomföra analyser på aggregerad och/eller individnivå för att kategorisera kunder inför riktade erbjudanden baserade på bland annat geodata (COOP, AXFOOD, ICA) och köphistorik (COOP, AXFOOD, ICA, ORGX) vilket innebär att olika kunder kan få olika erbjudanden via medlemskommunikation oavsett preferenser rörande fysiska butiker och e-handeln. Just e-handeln är dock där effekten av dataanvändning är som mest synlig för kund, där de produkter som visas baseras på köphistorik samt hur webbsidan navigerats tidigare (COOP, AXFOOD, ICA, ORGX).

Respondenten på Axfood beskriver att de inte har några avancerade maskininlärningsmetoder, men att de i begränsad omfattning använder klustermetoder för klassificering av kundsegment för att bli mer relevanta i sin marknadsföring. Hos ORGX har denna process till viss del automatiserats och respondenten berättar om hur de, genom maskininläring, utvecklat ett system som föreslår varor som passar det som redan finns i varukorgen eller vad kunden historiskt har handlat. Systemet skickar i tillägg meddelanden om outnyttjade mängdrabatter och respondenten beskriver det som mycket framgångsrikt och lönsamt. Övriga aktörer påvisar ingen liknande användning i nuläget, utan erbjuder främst viss recept-integration, dvs att ingredienser kan överföras från recept till varukorg.

4.3.2 Planering och utveckling

Samtliga aktörer använder insamlad persondata för att planera och utveckla sina respektive verksamheter. Användargenererad data som klick och besökshistorik, hur kund navigerat hemsidan och hur länge används för att optimera kundupplevelsen online och t ex geodata används för att optimera lokala erbjudanden eller planera nya butiker och lager (COOP, AXFOOD, ICA, ORGX). Persondata används i tillägg för att utvärdera marknadsföring och ICA använder både AI och maskininläring för analys av centrala strategier, om än i begränsad utsträckning. COOP analyserar Big Data för insikter om prisjusteringar, men det är oklart hur avancerad analysen är. Respondenten på ORGX beskriver hur data varit av stor vikt i framtagandet av nya automatiserade lager. Respektive lager ingår i ett spoke-hub nätverk där order behandlas i ett större lager (hub), och därefter förflyttas till mindre magasin (spokes) vilket möjliggör effektiv distribution av varor.

4.3.3 Kundkontakt

Aktörerna på den svenska marknaden lagrar och använder persondata i huvudsak för att underlätta vid reklamationssärenden och andra kontakter mellan kund och företag, till exempel teknisk support. ORGX använder sig i tillägg av maskininlärning för att kategorisera och ge olika prioritet för inkommande mail. Algoritmen är utvecklad med TensorFlow⁴ och bygger på den grundläggande frågeställningen: vad vill kunden? Respondenten på ORGX berättar om hur undersökningar visade att över 5% av alla mailkontakter saknade behov för svar, och hur de genom att träna systemet med stora mängder data och olika faktorer som används för att bestämma prioritet, lyckats sortera bort sådana mail automatiskt. Vidare tilldelas mail prioritet och fördelas per automatik, vilket inneburit ett uppsving i kundbemötande då mer tid kan ägnas till prioriterade ärenden snarare än att sortera mail.

4.3.4 Säkerhet

Säkerhetsarbetet är en viktig del av aktörernas arbete med både dataskydd och e-handel, varpå en majoritet (AXFOOD, ICA, ORGX) använder persondata för att säkerställa att t ex webbplats och självscanning används korrekt och av välmenade syften. För att minimera risken för säkerhetsbrister och manipulation erbjuds ingen detaljerad information om hur det praktiskt fungerar överlag med undantag för ORGX. ORGX använder maskininlärning för att analysera olika mönster för att avgöra om det rör sig om bedrägerier eller legitima köp. Algoritmen har utvecklats med TensorFlow och Google Cloud, och försöker urskilja de kunder som lagt in felaktiga uppgifter eller utgångna kreditkort av misstag från de som gör det av bedrägliga skäl, och baseras på en lång rad faktorer hämtade från tidigare köp. Satsningen har varit lyckad och respondenten beskriver hur de upptäcker bedrägerier 15 gånger mer effektivt än tidigare, vilket medfört kostnadsbesparingar både i kundtjänst och förlorade intäkter.

4.4 Framtidssyn

Samtliga av aktörerna på den svenska marknaden beskriver hur e-handeln blir allt viktigare och att de på olika vis rullar ut sina respektive satsningar inför framtiden, men också hur de tror att den fysiska butiken kommer vara viktig en lång tid framöver. Aktörernas olika

⁴ TensorFlow är ett mjukvarubibliotek utvecklat av Google Brain Team. Tillgänglig mjukvara är open-source och har starkt stöd för användning inom bland annat maskininlärning (TensorFlow 2018)

satsningar på e-handel har hittills handlat om att skapa en förlängning av den fysiska butiken genom att erbjuda samma sortiment, erbjudanden och priser online.

Axfoods ägare Axel Johansson är mycket intresserad av att utveckla arbetet kring Big Data. Respondenten på Axfood anser därför att det med stor sannolikhet kommer hända mycket på området framöver. Åkesson på hemköp beskriver hur de i framtiden kommer fortsätta arbeta med att integrera e-handeln med de fysiska butikerna så att de kompletterar varandra naturligt. I dagsläget består hemköps e-handel främst av "bunkringsvaror" och Åkesson menar att ett viktigt utvecklingsområde är att skapa merförsäljningsvaror i e-handeln, d.v.s. öka storleken på kundernas ordrar.

COOPs Claes Hassel tror att kombinationen av fysisk butik och e-handel är det som kommer att vara vinnande i längden. I nuläget sker merparten av packningen för COOPs e-handel i respektive butik, men Hassel beskriver även hur de etablerat särskilda plocklager i de fall omsättningen online vuxit till den grad att den stör den ordinarie verksamheten. Han tror även att kundfokus kommer att bli viktigare och förutspår att alla aktörer inom tre år kommer ha ett omnierbudande, dvs både fysisk butik och e-handel, som man utgår från kunden. Vidare beskriver Hassel hur han tror att den viktigaste utvecklingen av e-handeln är rörande hur kundgränssnittet kommer se ut, och nämner röstgränssnitt⁵ som en intressant framtida beställningsmetod.

För ICA handlar mycket av framtidsvisionerna kring e-handel om logistiken. ICAs e-handelschef Magnus Rosén anser att om e-handels volymerna ökar är det inte otänkbart att ICA etablerar så kallade dark stores, distributionscenter som enbart hanterar e-handel. Införandet av dark stores skulle förändra ICA:s logistikarbete radikalt men kan vara nödvändigt för lönsam e-handel av stor volym. Rosén beskriver hur butikens roll kan komma att utvecklas där de rutinmässiga större inköpen kommer genomföras online medan de fysiska butikerna kommer erbjuda andra områden som t ex färdigmat, inspiration eller upplevelser. Vidare tror han att E-handeln kommer förändra ytbehovet, dvs antal fysiska butiker och storlek på dem, när större andel av handeln sker i en onlinebutik. Rosén tror även att gränssnitt baserade på röstkommandon kan komma att få stor vikt i framtiden.

⁵ Röstgränssnitt innebär att röstkommandon kan användas för att styra olika funktioner, exempelvis genom en telefon eller personliga assistenter som Amazon Alexa.

Arbetet kring ICA:s digitala utveckling leds av innovationshuben ICAX som både utvecklar egna tekniska lösningar samt undersöker eventuella partnerskap. För framtiden undersöks bland annat automatisering och hur data kan användas för yt- och sortimentoptimering⁶, och ICA har under 2018 ingått ett avtal med brittiska Ocado där 600 miljoner investeras mellan 2018 och 2022 i en helautomatiserad e-handelslösning (ICA, 2018b). Syftet är att lägga grunden för nästa generations e-handelserbjudande, och skapa branschens mest effektiva lösningar även online (ibid).

ORGX har sedan lång tid haft stort fokus på teknik och tror fortsatt starkt på potentialen i maskininlärning när det gäller e-handeln. Respondenten beskriver en vision om en framtid där man kan nå så pass avancerade analyser att system kan förutspå vad en kund behöver utan att de tänkt tanken själva. I en närmare framtid beskrivs dock analyser av det egna systemet som av större vikt, med fokus på att upptäcka problem och påbörja åtgärd innan kunden upptäckt felet. Arbete med denna teknik har redan påbörjats, men respondenten beskriver hur det är en bra bit kvar innan det fungerar optimalt.

4.5 Utmaningar

Samtliga aktörer beskriver vilka möjligheter som följer med e-handeln, men beskriver också ett flertal olika utmaningar i form utav kundbeteenden och logistik. Hemköps Anna Åkesson beskriver till exempel hur kunder har olika köpbeteenden i butik och online, där fysiska butiker till stor del för kompletteringsköp med låg planeringsfront och online-handeln mestadels består av bunkring av basvaror. Detta medför en utmaning i att ta fram merförsäljningsvaror. Övriga utvecklingsområden som Åkesson ser är logistikkedjan både för onlinebutiker och fysiska butiker eftersom tre olika temperaturzoner krävs, och hon resonerar vidare kring olika lösningar i form av pickup-points, drönare eller lösningar för det uppkopplade hemmet. Åkesson beskriver vidare hur e-handeln och den tekniska utvecklingen medför en mängd olika möjligheter till individualisering, men anser samtidigt att *“...det måste finnas en balans mellan vad som är relevant och vad som är creepy”*.

⁶ Ytoptimering innebär att optimera hur varor placeras dels för att maximera kundexponering men även för att maximera hur mycket varor som får plats i butik eller lager. Sortimentoptimering innebär att orderstorlekar och val av varor anpassas för att möta kundernas efterfrågan.

COOPs Claes Hassel jämför övergången till e-handel med övergången från lanthandel till stormarknad vilken hade stor effekt på hur kundkommunikationen och kundhanteringen bedrevs, och att en omställning av denna kaliber kommer att innebära stora utmaningar. Köpbeteendet hos kunderna ändras med tekniken och Hassel exemplifierar detta genom vad han kallar matkassebeteendet, vilket tidigare inneburit 4–5 måltider per handling med färdigkomponerade recept. På senare tid har handling mer börjat ske på receptbasis där snittmåltider per handling ligger på 2.7 vilket Hassel tror beror på att folk vill kunna vara mer flexibla genom att handla mindre och oftare. Hassel lyfter i tillägg “last mile”, transporten mellan butik/lager och kund, som en viktig logistisk möjlighet och utmaning som e-handlarna hittills har glömt bort.

För ICA är kundbeteende en av de stora frågorna och för att lära sig av en mer utvecklad marknad beskriver Rosén hur ICA hämtar inspiration från Storbritannien och dagligvaruhuskedjan Tesco. Fokus har legat på att se hur de löst de logistiska utmaningarna såsom korta leveranser och hög flexibilitet. Rosén beskriver hur det nuvarande köpbeteendet är fokuserat på veckohandel och att utvecklingen går mot ett allt bredare sortiment eftersom e-handeln till stor del bygger på bekvämlighet. Detta skapar utmaningar eftersom ICA:s nuvarande e-handel utgår från butikernas respektive utbud. Rosén menar att anledningen till att inte fler handlare anslutit sig grundar sig i att ICA-butiker i regel samlar flera verksamheter och att handlarna, utifrån lokala möjligheter och intressen, ofta väljer att fokusera på ett område i taget.

ORGX ser inte kundbeteende som en utmaning då de kan hantera det efter så många års erfarenhet och istället är det rörande arbetet med GaaS som respondenter beskriver utmaningar. Utmaningarna är främst relaterade till att förmedla vilket värde som en investering i e-handel kan medföra och att det är stor skillnad på att sälja online jämfört med fysisk butik. Respondenten på ORGX beskriver hur artiklar som understryker vikten av e-handel har varit till viss hjälp för att övertyga potentiella kunder, men också att det är fortsatt utmanande att etablera nya kunder eftersom teknik ofta ses som någonting nytt och främmande. En ytterligare utmaning är att hantera flera olika kundbehov, dels B2B med GaaS och dels B2C med vanlig e-handel, med endast ett ingenjörsteam. Vidare beskriver respondenter behovet av att förtydliga affärsmodellen, då många av de nuvarande avtalen kring GaaS baseras på antalet ordrar medan ORGX samtidigt utvecklar maskininlärning applikationer som kraftigt

ökar varukorsstorlek, vilket gynnar kunden (B2B) men inte ORGX. Byggandet av relationer som gynnar båda parter är därmed en av organisationens stora utmaningar.

“Even if we have success stories we still have struggles creating a data culture where the value of data is understood within all areas of the company.” - respondent, ORGX

Samtidigt som organisationen kommit långt i arbetet med Big Data och har flera framgångsberättelser beskriver respondenten en svårighet med att etablera en datakultur där värdet av data förstås och uppskattas. Liknande utmaningar beskrivs i arbetet med t ex maskininlärning, där det alltid är svårt att uppskatta värdet av en applikation innan den har testats och därmed också svårt att få klartecken för ett nytt projekt. Som exempel ges att det tog 6 månader att få klartecken för applikationen som föreslår varor precis innan kassan, men det påpekas samtidigt att det gick mycket fort för version 2 av samma applikation eftersom potentialen redan bevisats. Vidare beskrivs GDPR som en ytterst aktuell utmaning i form av att anonymisera data och förändrade metoder för datainsamling. GDPR beskrivs vidare även som en utmaning eftersom det är oklart vilken typ av data som är användbar i framtiden.

5 Analys

5.1 Datakvalitet

Samtliga aktörer samlar in och använder liknande data i form av köp- och transaktionshistorik, kundfeedback och webbanvändningsdata. Med avseende på de kriterier som Clarke (2015) beskriver anser vi att köp- och transaktionshistorik är data som uppfyller D5-, D6- och D7-kriterierna, då de tydligt specificerar när transaktioner skedde, vad som inhandlades, var varorna beställdes ifrån samt vem som köpte dem. Denna data är starkt knutna till produkter, om syftet är att mäta och förutspå efterfrågan på olika produkter uppfyller dessa datakriterier, D1, D2, D3 och D4. För detta syftet anser vi därmed att denna data har hög kvalitet. Om syftet är att modellera efterfrågan hos olika kundgrupper måste även data om köparna användas.

Kundernas kontaktuppgifter, t.ex. bostadsområde och information angiven på medlemsidorna, t.ex. ålder och kön, är starkt knutna till kunderna, och uppfyller därmed D2 och D3 (Clarke, 2015). Denna data ger dock inte en fullständig bild av den komplexa individ som köper mat. Därför anser vi att denna data inte uppfyller D7. Slutsatsen kring denna data blir att den är pålitlig i den mening att den är exakt och starkt knuten till kunden, men den ger samtidigt inte en komplett bild av kunden och kan heller inte fullständigt förklara kundens köpbeteende. Eventuella korrelationer med denna data och köpmönster bör inte nödvändigtvis vara kausala då det kan finnas många dolda variabler som påverkar kunderna.

Data från kundservicekorrespondens kan ge mycket information om kunden, men den är till stor del ostrukturerad och svår att analysera. Ett möjligt användningsområde för denna data är som träningsdata för kundservicebotar som antingen kan ge svar på enklare frågor, eller, som för ORGX, sortera inkommande mail automatiskt genom att avgöra prioritet. Seth Grimes (2008) påstår att 80% av affärsrelevant är data ostrukturerad. Därför tror vi att det generellt kan vara värt för dagligvaruhandlarna att lägga resurser i sitt strategiarbete på att analysera hur de kan utnyttja ostrukturerad data bättre.

För att kunna utnyttja potentialen från Big Data till fullo tror vi att mer allmän och extern data bör samlas in för att korreleras med persondata och intern data. Dessa data kan innefatta väderdata, data angående viktiga händelser i lokalområdet eller data om sociala trender. Dels

för att kunna skapa Big Data nudges som Carolan (2018) beskriver, men även för att ge ett bredare underlag för prediktiva algoritmer, exempelvis vid efterfrågansprediktion. Vid hantering av mer abstrakt data där sambanden med det som ska beskrivas eller förutspås inte är uppenbart tycker vi att stor försiktighet bör iakttas. Denna abstrakta data kan skapa mycket värde i form av insikter som människor aldrig skulle ha hittat, men även här är det viktigt att inte dra förhastade slutsatser om kausalitet (Marr, 2015c). Insikterna bör granskas av någon med domänkunskap⁷ för att säkerställa att de är rimliga.

Sammanfattningsvis anser vi att den största andelen data som används är av god kvalitet. För en fullständigt datadriven affärsmodell tror vi dock att de svenska dagligvaruhandlarna måste vidga spektrumet av datakällor och börja utnyttja mer ostrukturerad data.

5.2 Dataanvändning

De olika aktörerna beskriver hur de huvudsakligen använder data för att förbättra upplevelsen i användandet av de tjänster som respektive aktör erbjuder, men också den egna lönsamheten. Samtliga aktörer beskriver dataanvändning som syftar till att skapa bättre förståelse för kunder och deras beteenden vilket sedan påverkar till exempel marknadsföring i form av riktade kampanjer och butiksutveckling i form av priser, sortiment och läge av t ex butik eller lager. Användande av data för att skapa större förståelse för kunder är ett vanligt användningsområde för Big Data (Bose, 2009; Marr, 2015b; Hitchcock, 2017), men det är även tydligt att utvecklingen kommit olika långt för de olika aktörerna. Den individanpassning som de svenska aktörerna beskriver i form av riktad marknadsföring och webbanpassning liknar vad Thaler (2008) beskriver som nudges, men det är endast ORGX som tydligt arbetar med datadrivna nudges. De exemplifieras bäst genom systemet som, baserat på maskininlärning, påminner vid kassan om produkter som kunden kanske glömt vilket är ett tydligt exempel på hur Big Data kan användas. ORGX användning av Big Data för att skapa optimerade nudges har enligt respondenten inneburit mer relevanta personliga erbjudanden och ökad korgstorlek, vilket väl stämmer överens med den potential i nudges som Carolan (2018) beskriver.

Utöver kundinsikter är ett annat vanligt användningsområde för Big Data att utgöra en grund för beslutsfattande (Bose, 2009; McAfee & Brynjolfsson, 2012; Marr, 2015; Hitchcock,

⁷ Domänkunskap innebär kunskap om ett specifikt område, ofta den miljö ett specifikt system verkar i.

2017). Även om ICA i viss mån använder AI för stöd i det strategiska arbetet är det endast ORGX som uppger att Big Data genomsyrar verksamheten och används till grund för beslut i större utsträckning. Det är också bara hos ORGX som vi funnit datadriven processoptimering internt i form av automatiserade lager, automatiserad mailsortering och AI-baserat bedrägeriskydd. Hos de svenska aktörerna har vi inte funnit någon liknande användning, men vissa paralleller kan dras till att data används för att utveckla verksamheten. I empirin framgår dock inte vilken tyngd som data har när det rör beslut inom detta område, och det kan mycket väl vara så att data bara används i begränsad omfattning. Vidare uppger de svenska aktörerna att data lagras och används av säkerhetsskäl, men varken respondenten från Axfood eller något annat empiriskt underlag pekar på att det rör sig om ett AI-baserat aktivt skydd. Istället rör det sig troligen om att skapa ett underlag om någonting skulle ske, snarare än att förhindra att det sker. Det kan dock vara så att beskrivning av säkerhetsarbetet medvetet utelämnats för att försvåra för bedragare att hitta öppningar, vilket gör det svårt att dra någon egentlig slutsats.

Överlag verkar det finnas skillnader mellan de svenska aktörerna och den brittiska i hur de använder data, trots att datainsamling och datakvalitet verkar vara relativt snarlika. Detta kan bero på att företag generellt inte kommit långt i användandet av data (DalleMule & Davenport, 2017) men också på att en stor del av värdefulla insikter finns i ostrukturerad data (Grimes, 2008) vilken är både svårare och dyrare att analysera än strukturerad data (Marr, 2015c). Ett viktigt perspektiv är dock att medan ORGX aktivt arbetar med Big Data, så kan det vara så att de svenska aktörerna i nuläget arbetar med data fränkopplat Big Data som begrepp. Det skulle i sådana fall kunna innebära att den data som de samlas in bara är ett nytt medium för uppgifter som de alltid använt, och att det överlag saknas kompetens för att dra nytta ur Big Data. Skillnaden kan också bero på en ovilja att göra nödvändiga investeringar eftersom värdet inte är tydligt, vilket även respondenten från ORGX menar är vanligt förekommande för nya projekt.

5.3 Mognad

Aktörernas användande av data för att segmentera kundgrupper, skapa prognoser och anpassa sortiment tyder på att de åtminstone nått steg två i Schmarzos (2015) BDBMMI-modell, *business insights*, där fokus ligger på att ta fram värdefulla insikter ur stora datamängder. Vi ser dock vissa indikationer på mer avancerad användning i form av att varor och

marknadsföring anpassas efter tidigare val vilket tyder på användandet av nudges (Thaler, 2008), och ICA beskriver hur de påbörjat arbetet med att använda maskininlärning för att optimera priser och marknadsföring. Det innebär att det åtminstone för ICA verkar finnas indikationer på det tredje steget i Schmarzos (2015) BDBMMI-modell, *business optimization*, eftersom de använder data för att optimera verksamheten centralt. Användandet av maskininlärning för marknadsföring kan i tillägg tyda på att försäljningsdata korreleras med tidigare utskick till kund, vilket på sikt kan innebära att optimerade nudges kan komma att genereras automatiskt liksom Carolan (2018) beskriver.

För övriga svenska aktörer är det inte lika tydligt huruvida de använder sig av datadriven optimering eller inte, och i sådana fall på vilken nivå. Bolagen inom Axfood uppger förvisso att data används för personliga erbjudanden och att utveckla verksamheten vilket kan tyda på steg tre (Schmarzo, 2015), men respondenten vi intervjuade menade samtidigt att de inte har några modeller för maskininlärning vilket kan tyda på att utvecklingen inte kommit särskilt långt. Detsamma tycks gälla för COOP som förvisso analyserar big data, men där vi i övrigt inte funnit någon särskild eller avvikande användning. ORGX anser vi ligger på den sista nivån, *business metamorphosis (ibid)*, trots att de inte har omvandlat sin verksamhet. ORGX startade nämligen sin verksamhet med en affärsmodell som de svenska aktörerna kan transformera till, och hela ORGX verksamhet genomsyras av det digitala, datadrivna och automatiserade där data används såväl till beslut som för att generera nya intäkter. Systemet som föreslår varor som kunden kanske glömt anser vi är ett tydligt exempel på de Big Data nudges som Carolan (2018) beskriver eftersom de genereras automatiskt genom maskininlärning. Systemen för att förhindra bedrägerier och automatiskt sortera mail är i sin tur tydliga exempel på hur Big Data används i praktiken, vilket vi anser bekräftar den potential i dataanvändning som Marr (2015) och Hitchcock (2017) föreslår.

Sammantaget så tycks det inte råda några större skillnader i anledning till dataanvändning, utan det är i tekniskt avancemang och integration som vi ser de stora skillnaderna mellan svenska aktörer och den brittiska där den senare kommit avsevärt mycket längre i sin utveckling. Detta exemplifieras bland annat i form av användandet av maskininlärning för att optimera vissa delar av verksamheten eller för att driva nudges medan svenska aktörer på sin höjd verkar använda kundsegmentering och viss individualisering. Svenska aktörer tycks alltså vara efter i användandet av data, trots god tillgång och kvalitet på den, och frågan är varför det ligger till på det sättet. ORGX startade sin verksamhet i början av millennieskiftet

vilket var ungefär samtidigt som de svenska aktörerna genomförde sina första storsatsningar på e-handel, men med väldigt olika slutresultat. Detta väcker tankar om att det eventuellt kan röra sig om olika kundmönster eller olika kultur, där den svenska marknaden mognat betydligt senare än den brittiska. En annan tänkbar anledning som vi ser är att respektive marknad börjat i olika ändar av den SMART-modell som Marr (2015a) presenterat, vilken understryker vikten av att börja med strategi snarare än att försöka transformera på en gång. Om de svenska aktörerna började transformera endast för att alla andra gjorde det eller för att det var på modet, utan att grunda det i den övergripande strategin, så kan det vara en förklaring.

5.4 Vägen framåt

Som vi har sett är de svenska aktörerna i ett tidigt skede med avseende på utnyttjandet av Big Data, men verkar ha inställningen att utvecklas inom detta område och gärna snabbt. Nedan beskrivs och diskuteras framtida utmaningar som vi har identifierat samt möjliga sätt att hantera dessa.

5.4.1 Syfte och mål med dataanvändningen

Utan en tydlig plan och strategi riskerar övergången till ökad användning av Big Data bli mycket kostsam och överksam. Vi anser att det är viktigt att inte planlöst anställa data scientists och tro att de ska lösa alla problem. Som Marazzi (2017) beskriver är det viktigt att ledningen har förståelse för Big Data för att kunna ge rimliga arbetsförutsättningar för data scientists. Detta verkar vara en utbredd utmaning då även respondenten i ORGX nämnde det som ett problem. Även i ORGX där affärsmodellen är utformad för att vara datadriven var det problematiskt att motivera värdet av data för ledningen. I Marrs (2015a) SMART modell är det första steget *Start with Strategy* vilket innefattar att ledningen förstår innebörden av Big Data och kopplingen till strategi. En tydlig strategi är även viktig för att på förhand klargöra vilka typer av insikter man letar efter och vilka typer av optimeringar eller omställningar som måste göras.

Angående M:et i SMART, *Measure and Metrics*, finns det som ovan nämnt antagligen utrymme för de svenska aktörerna att utnyttja mer extern data. Antingen kan de samla in data själva eller så kan de köpa data. Samtidigt är den främsta poängen i *Measure and Metrics*, att företagen måste vara noggranna med att definiera syftet med datainsamlingen. Detta är viktigt

dels för att följa GDPR, men också för att slippa lägga pengar och energi på data som inte har något värde.

Både i SMART modellen (Marr, 2015a) och i BDBMMI-modellen (Schmarzo, 2015) handlar det avslutande steget om transformation av affärsmodellerna, *Transform your business* i SMART och *Business Metamorphosis* i BDBMMI. Det är tydligt att alla studerade aktörer tror på värdet i Big Data, men de svenska aktörerna verkar främst fokuserade på att förstärka sin nuvarande affärsmodell med Big Data. ICA har en stor marknadsandel och visar tydligast satsning på teknik, dock verkar deras visioner endast sträcka sig till Business Optimization fasen i BDBMMI. För de mindre aktörerna kanske det kan vara läge att utvärdera möjligheterna för data monetization eller att helt transformera affärsmodellen för att kunna disruptera marknaden för att kunna konkurrera med ICA.

5.4.2 Kompetens

En del som vi anser som viktig i data strategin är vilken kompetens som ska arbeta med det. När *Apply analytics* och *Report results* stegen i SMART nås måste rätt personal finnas tillgänglig. Att använda data-tolkare som Ariker et al. (2014) beskriver är enligt oss extra viktigt för företagen i dagligvarubranschen som traditionellt sett inte har arbetat mycket med Big Data- och AI-teknologier. Att ha anställda som kan fungera som en brygga mellan ledningen och data-teamet tror vi kan göra kommunikationen enklare och kan hjälpa till att skapa ömsesidig förståelse för båda sidorna. Detta för att slippa de problem som Marazzi (2017) beskriver angående en ledning som bara förväntar sig resultat från data-teamet utan att tillhandahålla rätt material och utan att ställa rätt frågor.

En fråga som företagen även behöver ställa sig är vad de ska utveckla själva och vilka tjänster som ska köpas. I fallen där AI ska användas för automatisering och effektivisering, t.ex. automatiska lager, röstgränssnitt eller chattbotar i kundtjänst, är det enligt oss en god idé att köpa tjänster då dessa uppgifter är relativt standardiserade, en chattbot på ICA fungerar ungefär likadant som en chattbot på COOP. Då det finns företag, som är experter på att ta fram dessa system är det antagligen lättare och billigare att köpa dessa tjänster. ICAs samarbete med Ocado kring automatiserade lager är ett exempel på just detta, där ICA utnyttjar Ocados färdiga lösning för automatiserade lager för att snabbt kunna implementera det.

För uppgifterna relaterade till data-mining anser vi att det är bra att ha en egen avdelning. Marr (2015c) nämner att en av de största riskerna med Big Data är dålig analys. För att undvika att ta felaktiga beslut från data mining insikter och för att kunna ställa rätt frågor tror vi att det är viktigt med djup insikt i den specifika branschen och det specifika företaget. Detta går i linje med idén om data-tolkare som Ariker et al. (2014) beskriver. Att skaffa denna domänkunskap tror vi är lättast om personerna arbetar i den egna organisationen, då detta ger tid att lära känna och förstå organisationen på djupet.

5.4.3 GDPR

En av de stora möjligheterna med Big Data som nämns både av de svenska aktörerna och teorin är förmågan att kunna individualisera kunderbjudande och kundbemötande. Detta kräver dock att personlig data om varje individ samlas in. GDPR tvingar företag till transparens kring sitt dataanvändande och ger även alla individer rätten till att bli bortglömd, d.v.s. få sin data raderad (Datainspektionen, n.d). Förutsatt att många individer väljer att få sin data borttagen begränsas givetvis företagets förmåga att använda Big data för att individualisera sina erbjudanden. Det är alltså viktigt hur företag kommunicerar kring sitt användande av personuppgifter. God kommunikation och verklig transparens tillåter företagen att samla in värdefull persondata med individernas tillåtelse. Även här tror vi att det är viktigt att företagen tydligt definierar för sig själva i vilket syfte de insamlar data. Marr (2015a, 2015c) skriver om vikten av att endast samla in data som har ett tydligt syfte då det kan vara dyrt att samla in för mycket data. Om det inte finns ett tydligt syfte med att samla in personuppgifter tas en stor risk utan att något värde tillförs. Vi tror även att det är lättare att kommunicera ut ett syfte med dataanvändningen om syftet är tydligt definierat internt.

5.4.4 Logistik

De flesta aktörerna nämner logistiken som en stor nuvarande och framtida utmaning allt eftersom e-handeln går framåt. Empirin talar för last mile som en stor utmaning, men också hur e-handeln i vissa områden blivit så stor att den stör den ordinarie verksamheten med den nuvarande modellen att plock sker i butik. Som lösning på dessa utmaningar har COOP i nuläget etablerat en hybridlösning som innebär en centraliserad lagerlösning, ICA studerar Tesco och diskuterar så kallade dark stores, och Axfood har valt att ta hjälp av nischade aktörer genom förvärven av mat.se och middagsfrid.

Vilken typ av logistiklösning som är bäst kan endast framtiden avgöra och vi tror att det kommer bero mycket på hur kundbeteendet och e-handeln i sig utvecklas. Eventuellt kan de

svenska aktörerna ta lärdom av hur ORGX arbetar med spoke-hub nätverk och automatiserade lager, men det kan också vara så att den typen av lösning endast fungerar kring storstäder och inte är lämplig som en riksomfattande lösning. ICA kan troligen lära sig mycket genom samarbetet med Ocado och deras automatiserade lagerlösning, och om investeringen blir lyckad tror vi att det är sannolikt att fler aktörer anammar ett liknande upplägg.

Sammanfattningsvis är det mycket som håller på att hända i dagligvarubranschen och många av de svenska företagen verkar vara precis i uppstartsfasen av sin digitala resa. Då flera internationella företag har kommit långt finns det goda möjligheter för de svenska företagen att lära sig av andra, men det är likväl viktigt att de grundar sina beslut i strategi och inte blir förblindade av alla möjligheter. I tillägg är vår bedömning att den övergripande tillgången på data, samt kvaliteten på densamma, är god vilket ger en goda förutsättningar för att utvecklas snabbt tekniskt. Det återstår dock att se om framtidens dagligvaruhandel innebär leverans med drönare och beställning via AI, men möjligheterna finns där redan idag. Samtidigt är frågan vart gränsen egentligen går för när det bekväma och personliga övergår till obehagligt, och det är möjligt att det blir en allt viktigare fråga allt eftersom tekniken utvecklas.

6 Slutsatser

Hur används Big Data i dagligvaruhandeln i nuläget?

Vad som framkommer i studien är att det förekommer skillnader i vilken omfattning som Big Data används i den svenska dagligvaruhandeln jämfört med den brittiska aktören. De svenska aktörerna ser ut att vara ganska unga i sitt arbete med Big Data och ligger kring nivå två eller tre i Schmarzos (2015) BDBMMI-modell vilket innebär att de i viss mån utnyttjar data för att ta fram insikter om marknadsläget och sina kunder. De har även i viss utsträckning börjat använda data för att optimera sin verksamhet i form av sortiment, priser och personliga kunderbjudanden. Den brittiske aktören bedöms i sin tur befinna sig på den sista nivån vilket innebär att hela verksamheten transformerats och optimerats med hjälp av Big Data. Big Data lägger grunden för en stor del av deras beslut, och artificiell intelligens används för att skapa merförsäljning, avlasta kundtjänst samt förhindra bedrägerier online.

Vilka risker och möjligheter finns med Big Data?

Big Data för med sig möjligheter främst i form av att kunna mäta och kvantifiera i princip allt vilket innebär att företag kan nå helt nya insikter om sin verksamhet. Studien identifierar möjligheter för Big Data inom efterfrågansprognoser och prisoptimering vilka kan användas för att optimera verksamheten, men också som beslutsunderlag i strategifrågor. Studien ser i tillägg möjligheter inom effektiv marknadsföring, individualisering och automatisering som drivs av Big Data.

Samtidigt som Big Data för med sig ett antal möjligheter så finns det även ett antal risker. Studien identifierar en risk i form av att värdet i satsningar på Big Data inte alltid är tydligt, särskilt om det rör en ny applikation, och i tillägg identifieras svårigheter i form av att engagera ledningen i data-relaterade frågor. GDPR behöver inte nödvändigtvis vara ett hinder, utan kan snarare vara något positivt då företagen tvingas att definiera ett tydligt syfte med data som samlas in. Detta kan hjälpa dem till ett effektivt användande av Big Data. Vid uppmärksamhet och transparens angående hur data används bör inte GDPR utgöra något större hot för företagen.

Hur kan användningen av Big Data se ut i framtiden?

Studien visar att de studerade svenska aktörerna är intresserade av att expandera sin e-handel och i många fall utveckla sitt dataarbete. Vidare så använder aktörerna liknande datakällor, data bedöms överlag hålla bra kvalitet och riskerna som hanterbara vilket innebär att de svenska företagen har en stabil grund för att utveckla sitt arbete med Big Data.

Studien indikerar att användningen av Big Data för den svenska marknaden troligen kommer att röra sig mot ett ökat fokus på optimering och individanpassning, men inte till vilken utsträckning. Studien lyfter dock att utvecklingen kan komma att accelereras genom partnerskap med utländska aktörer, och att det kan innebära att maskininlärning likt den hos den brittiske aktören kan bli aktuell. Vidare anses att värdet i att använda mer abstrakt och ostrukturerad data bör diskuteras. Oavsett vilket spår som utvecklingen tar så betonar studien vikten av ett kritiskt förhållningssätt med en tydlig strategi och plan för hur utvecklingen ska gå till.

7 Vidare forskning

Denna rapport har fokuserat på de övergripande strategiska möjligheterna och riskerna med Big Data, specifikt inom den svenska dagligvaruhandeln. Det som upptäcktes var att de stora svenska aktörerna till stor del står inför en stor förändringsprocess. En möjlig framtida forskning är att utifrån ett change management perspektiv studera hur själva förändringsprocessen bör gå till.

En del resultat i denna studie går antagligen att generalisera till andra branscher än dagligvaruhandeln, men för ett djupare förståelse behöver troligen varje bransch studeras separat, vilket skulle kunna vara ett ämne för kommande forskning.

Då denna studie inte har fokuserat på specifika tekniska lösningar finns det mycket utrymme för fortsatt forskning angående hur specifika behov kan förverkligas tekniskt. Förslag på tekniska områden att undersöka:

- Om maskininlärning kan användas för att minska butikernas svinn.
- Analys av orderflöden för att ta fram optimala platser för distributionslager och darkstores.
- Automatisering av packning och leverans av e-handels ordrar.

Referenslista

Amazon (2013), *Method and system for anticipatory package shipping*, US 8615473

Ariker, Matt et al. (2014) *How to get the most from Big Data* Tillgänglig på:
<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/how-to-get-the-most-from-big-data>
Hämtad 2018-03-29

Axfood (2018) *Axfood i korthet* Tillgänglig på:
<https://www.axfood.se/om-axfood/axfood-i-korthet/>
Hämtad 2018-05-24

COOP (2018a) *Om COOP* Tillgänglig på:
<https://www.coop.se/globala-sidor/om-coop/>
Hämtad 2018-05-24

COOP (2018b) *Dina personuppgifter är viktiga för oss* Tillgänglig på:
<https://www.coop.se/Globala-sidor/Om-webbplatsen/personuppgifter/>
Hämtad 2018-05-24

BBC (2014) *Target data theft affected 70 million customers* Tillgänglig på:
<http://www.bbc.com/news/technology-25681013>
Hämtad 2018-05-22

Bose, R. (2009), *Advanced analytics: opportunities and challenges*. Industrial Management & Data Systems. Vol 109 Iss 2 pp. 155-172

Bryman, A. & Bell, E. (2011) *Företagsekonomiska forskningsmetoder*. 2:a utgåvan, Stockholm: Liber förlag.

Bughin, Jaques et al. (2018) *Why digital strategies fail* Tillgänglig på:
<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/why-digital-strategies-fail>
Hämtad 2018-05-14

Clarke, Roger (2015) *Big Data, big risks* Tillgänglig på:
<https://onlinelibrary-wiley-com.proxy.lib.chalmers.se/doi/full/10.1111/isj.12088>
Hämtad 2018-05-09

Coblin, Harvey et al. *Structured and Unstructured Data: What is It?* Tillgänglig på:
<https://sherpasoftware.com/blog/structured-and-unstructured-data-what-is-it/>
Hämtad 2018-05-10

Dahr, Vasant (2013) *Data Science and Prediction* Communications of the ACM Volume 56 Issue 12, Pages 64-73

DalleMule, Leandro och Davenport, Thomas H (2017) *What's Your Data Strategy?*

Tillgänglig på:

<https://hbr.org/2017/05/whats-your-data-strategy>

Hämtad 2018-04-02

Datainspektionen *Dataskyddsförordningen*

<https://www.datainspektionen.se/dataskyddsreformen/dataskyddsförordningen/>

Hämtad 2018-05-18

DIBS (2017) *SVENSK E-HANDEL* Tillgänglig på:

<http://info.dibs.se/hubfs/Svensk%20E-handel%202017-1.pdf?hsCtaTracking=2889946d-5879-449b-be87-b52e53a484a3%7C345c0a08-5aaf-4e9c-8e90-b29c3a819a3d>

Hämtad 2018-04-02

DLF, Delfi & HUI Research (2018) *Dagligvarukartan 2017* Tillgänglig på:

<http://www.hui.se/statistik-rapporter/index-och-barometrar/dagligvarukartan>

Hämtad: 2018-05-24

E-handelstrender (2017) 44. *Coop och Hemköp olika modeller för e-handel* Tillgänglig på:

<https://uppkopplat.se/category/ehandelstrender/>

Hämtad 2018-05-24

E-handelstrender (2018) 52. *ICA:s satsar fullt ut på e-handel* Tillgänglig på:

<https://uppkopplat.se/category/ehandelstrender/>

Hämtad 2018-05-24

Flasiński M. (2016) *Computational Intelligence. In: Introduction to Artificial Intelligence*. Springer, Cham

Grimes Seth (2008) *UNSTRUCTURED DATA AND THE 80 PERCENT RULE*, Tillgänglig på:

<https://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/>

Hämtad 2018-05-21

Gulla, Joseph (2012) *Seven Reasons IT Projects Fail* Tillgänglig på:

http://ibmsystemsmag.com/power/systems-management/workload-management/project_pitfalls/

Hämtad 2018-03-31

Hemköp *HEMKÖPS INTEGRITETSPOLICY* Tillgänglig på:

<https://www.hemkop.se/artikel/gdpr>

Hämtad 2018-05-24

Hitchcock, Erin (2017) *Five Big Data Use Cases for Retail* Tillgänglig på:

<https://www.datameer.com/company/datameer-blog/five-big-data-use-cases-retail/>

Hämtad 2018-03-29

ICA (2018a) : *ICA Gruppen i korthet* Tillgänglig på: <https://www.icagruppen.se/om-ica-gruppen/#!/>

Hämtad 2018-05-20

ICA (2018b) *ICA Gruppen tecknar avtal med Ocado om ny e-handelslösning*

Tillgänglig på: <https://www.icagruppen.se/arkiv/pressmeddelandearkiv/2018/ica-gruppen-tecknar-avtal-med-ocado-om-ny-e-handelslosning/>

Hämtad 2018-05-19

ICA (2018) *Vi skyddar dina personuppgifter* Tillgänglig på:

<https://www.ica.se/policies/behandling-av-personuppgifter/>

Hämtad 2018-05-24

ITU (2012) *Internet of Things Global Standards Initiative* Tillgänglig på:

<https://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx>

Hämtad 2018-05-24

Jacobsen, D. (2002) *Vad, hur och varför? - Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*. Lund: Studentlitteratur AB.

SvD (2017) *Storsatsar på e-handel – men går back miljoner* Tillgänglig på:

<https://www.svd.se/matjattar-storsatsar-pa-e-handel--forlorar-miljoner>

Hämtad 2018-04-02

Machiraju S., Modi R. (2018) *Conversations as Platforms*. In: *Developing Bots with Microsoft Bots Framework*. Apress, Berkeley, CA

McAfee, A. & Brynjolfsson, E. (2012) *Big Data: The Management Revolution*, Harvard Business Review October 2012, p. 59-68

Marazzi, Alan (2017) *The most difficult thing in data science: politics*.

<https://www.rdisorder.eu/2017/09/13/most-difficult-thing-data-science-politics/>

Hämtad 2018-05-10

Marr, Bernard (2015a) *Big data : using SMART Big Data, analytics and metrics to make better decisions and improve performance*, John Wiley & Sons

Marr, Bernard (2015b) *Big Data: A Game Changer In The Retail Sector*, Tillgänglig på:
<https://www.forbes.com/sites/bernardmarr/2015/11/10/big-data-a-game-changer-in-the-retail-sector/#496ef9f29f37>

Hämtad 2018-03-29

Marr, Bernard (2015c) *The Biggest Risks of Big Data*, Tillgänglig på:
<https://www.linkedin.com/pulse/5-biggest-risks-big-data-bernard-marr/>

Hämtad 2018-05-21

Marr, Bernard (2016) *Big Data: Big Data At Tesco: Real Time Analytics At The UK Grocery Retail Giant*, Tillgänglig på:

<https://www.forbes.com/sites/bernardmarr/2016/11/17/big-data-at-tesco-real-time-analytics-at-the-uk-grocery-retail-giant/#652025a161cf>

Miller, Paul (2012) *Tesco uses data for more than just loyalty cards*

Tillgänglig på http://cloudofdata.com/2012/10/tesco-uses-data-for-more-than-just-loyalty-cards/?utm_source=datafloq&utm_medium=ref&utm_campaign=datafloq

Schmarzo, Bill (2015) *Big Data MBA : Driving Business Strategies with Data Science*, John Wiley & Sons, Incorporated

Stockdale, R. & Standing, C. (2005) *An interpretative approach to evaluating information systems: A context, context, process framework*, European Journal of Operational Research 173(2006) p. 1090-1102

Svensk Digital Handel (2017) *Digital Mathandel* Tillgänglig på:

<https://dhandel.se/rapporter/digital-mathandel/>

Hämtad 2018-04-02

Thaler, Richard H. (2008) *Nudge: Improving Decisions About Health, Wealth, and Happiness* Yale University Press

Willys (2018) *Vår integritetspolicy* Tillgänglig på:

<https://www.willys.se/artikel/integritetspolicy>

Hämtad 2018-05-24

Bilaga 1: Frågemall

Allmänt

- Berätta övergripande om er verksamhet
- Vad är din roll i organisationen?
- När öppnade ni online, och av vilken anledning?
- Vilka är de största utmaningarna/möjligheterna för er verksamhet i dagsläget?
 - Hur kommer dessa förändras i framtiden?
 - Hur hanterar ni dem?

Big Data

- Vad är er syn på dataanalys och Big Data?
 - Vilka utmaningar/möjligheter ser ni med Big data?
 - Hur hanterar ni dem?
- Hur arbetar ni med dataanalys/Big Data i er organisation? (Övervakning/rapportering, Insikter, optimering, monetization skapa nya intäcksströmmar, helomställning/ servitization)
 - Vilka är det som arbetar med dataanalys i er organisation. (Data scientists, marknadsförare, analytiker)
 - Vilka insikter är intressanta? På vilket sätt?
 - Hur skapar dessa insikter värde?
 - Hur påverkas ert arbete av andra organisationer? Samarbeten? Inspiration osv.

Om de inte arbetar med big data eller dataanalys:

- Hur kommer det sig?
- Är det något som ni har övervägt?
 - På vilket sätt?

Datainsamling

- Hur samlar ni in data?
 - Vad är syftet med datainsamlingen?
 - Egen data eller branschorganisationer?
 - Vilken typ av data?
 - Hur stor andel data används?
 - Hur ser data ut? Strukturerad, ostrukturerad, realtid, batcher?
- Hur hanterar ni personuppgifter?
- Vad är dina tankar kring GDPR?
- Hur kommer GDPR påverka ert arbete med big data?

Strategi

- Hur ser strategin kring Big Data ut?
- Hur kopplas Big Data strategin till övergripande strategin?

Implementering/ förändring

- Hur gick implementeringen till? SMART
- Vilka förändringar krävdes? Tekniska/ organisatoriska?
- Vad blev resultatet? Värdering?

Avslutning

Hur ser er vision/drömscenario kring Big Data/dataanvändning ut?

