

Kritische succesfactoren voor de inzet van Big Data:

is groot wel anders?

Jan van den Belt en Charlotte Nillesen

SAMENVATTING De digitalisering van de maatschappij heeft een nieuw fenomeen voortgebracht: 'Big Data'. De onuitputtelijke bron "data" is de drijfveer van de economie. Het verkrijgen van deze data is niet het probleem, maar wel het omzetten van deze data in waardevolle informatie¹. Een toenemend aantal consumenten winkelt via websites en geeft haar mening over geleverde producten en diensten op daarvoor ingerichte internetsites of via sociale media zoals Twitter. Gesteld wordt dat wanneer een organisatie in staat is om deze gegevens op adequate wijze te verzamelen, te beheren en te analyseren, prestaties kunnen worden verbeterd door het nemen van betere beslissingen en het optimaliseren van processen.

Maar wat betekent het om zeer grote volumes gevarieerde data met hoge snelheid tijdig te transformeren naar nuttige informatie? Verschilt dat eigenlijk wel van de huidige werkwijzen met betrekking tot dataverwerking en informatiemanagement?

Door middel van een case study zijn bij zeven Nederlandse retailorganisaties beïnvloedende factoren voor implementatie van een big data-strategie bestudeerd. Hieruit blijkt dat een groot aantal succesfactoren voor 'traditionele' dataverwerking en informatiemanagement ook geldt voor big data, maar dat onder meer management support, bedrijfscultuur, definiëren en standaardiseren van data en geavanceerde data-analyse-expertise van extra kritisch belang zijn vanwege het afdelings- en systeemoverschrijdende karakter van de implementatie van een big data-strategie.

RELEVANTIE VOOR DE PRAKTIJK Veel onderzoek is te vinden over de relatie tussen informatie(overload) en de prestaties van een organisatie, als ook over informatiemanagementmethoden, zoals Datawarehousing, Business Intelligence en Enterprise Resource Systemen. Deze studie vergroot de kennis over factoren die al dan niet specifiek zijn voor de uitvoering van een big data-strategie, hetgeen tot uitdrukking komt in het 'Big Data Kritische Succesfactor Framework'.

Met name voor retailers lijkt door gebruik van big data de mogelijkheid te ontstaan om klanten beter te leren kennen en zelfs te kunnen sturen in hun koopgedrag,

zodat assortiment, inkoop en levering kunnen worden geoptimaliseerd. Als retailers in staat zijn om big data op een effectieve manier te gebruiken, kan het de mogelijkheid creëren om een concurrentievoordeel te behalen en te behouden.

1 Inleiding

1.1 Achtergrond

Veel traditionele data zijn door de toegenomen digitalisering van de maatschappij getransformeerd in zogenaamde 'Big Data' (Gartner, 2013; McAfee & Brynjolfsson, 2012; Minelli et al., 2013; Sathi, 2012; Zikopoulos et al., 2012). Sommigen zien big data als een hype², gestimuleerd door de IT-sector, anderen zijn van mening dat gebruik van big data noodzakelijk is voor organisaties om concurrerend te kunnen blijven. Feit is dat big data een overload aan informatie genereert (Chen, Chiang & Storey, 2012), hetgeen door organisaties moet worden opgelost.

In een studie met 586 senior executives, antwoordde bijna de helft dat strategisch (big) datamanagement een belangrijke factor is geweest bij het nemen van beslissingen in de afgelopen vijf jaar. Bijna 10% gaf aan dat data-analyse de wijze waarop de organisatie haar business runt compleet heeft veranderd (The Economist Intelligence Unit, 2011). Onderzoeksorganisatie Gartner schat dat organisaties die optimaal gebruik maken van big data en een adequate informatiestrategie hebben 20% beter zullen scoren in financieel opzicht dan hun concurrenten (Beyer, 2012). Gegeven het type klant lijkt dit bij uitstek van toepassing in de retailsector.

In dit artikel staat de volgende vraag centraal: *Wat zijn de kritische factoren voor een succesvolle implementatie van een big data-strategie in Nederlandse retailorganisaties?*

1.2 Aanpak onderzoek

Het onderzoek is exploratief van aard hetgeen betekent dat het zich richt op het verkrijgen van inzicht in de beginfase van een probleem en kan resulteren in een (kwantitatief) te toetsen hypothese (Eisenhardt, 1989). Een case study is een passende methode voor de beginfase van onderzoek over een bepaald onderwerp (Eisenhardt, 1989). Deze case study bestaat uit: een literatuurstudie, een online vragenlijst en diepte interviews (Staehr et al., 2012). Elf functionarissen verdeeld over zeven Nederlandse retailorganisaties hebben deelgenomen aan de case study. In de literatuurstudie is gefocust op wetenschappelijke onderzoeken die ingaan op kritische succesfactoren van informatiemanagementmethoden en -systemen in het algemeen. Daarnaast is specifiek gekeken naar de kenmerken van retailorganisaties en welke big data-toepassingen hierbij mogelijk zijn.

1.3 Opbouw van dit artikel

Allereerst wordt nader ingegaan op de begrippen big data en retailorganisatie. Daarna wordt een aantal kritische succesfactoren besproken die op basis van de literatuurstudie zijn gevonden en gelden voor verschillende informatiemanagementmethoden. Vervolgens worden de onderzoeksresultaten besproken met daarin de scores op de geïnventariseerde succesfactoren, alsmede nieuw genoemde succesfactoren bij iedere case, waarna de overeenkomsten en verschillen tussen de cases worden besproken. Tenslotte zal een “kritische succesfactor framework” voor Big Data worden gepresenteerd.

2 Theoretisch kader

2.1 Big Data

2.1.1 Van data naar kennis

Literatuur over informatie en kennis refereert vaak aan de data-informatie-kennis-wijsheid (DIKW)-hiërarchie, die het eerst werd geformuleerd door Ackoff. “*The implicit assumption is that data can be used to create information; information can be used to create knowledge, and knowledge can be used to create wisdom*” (Rowley, 2007, p. 164). Data is een product van dagelijkse werkzaamheden en observaties en is niet te gebruiken voordat het is getransformeerd in een voor mensen begrijpelijk formaat dat informatie wordt genoemd (Rowley, 2007; Laudon & Laudon, 2010). Nadat data in informatie is getransformeerd, kunnen gebruikers kennis en ervaring daaraan toevoegen (Rowley, 2007). Tenslotte wordt wijsheid gebruikt om de kennis toe te passen (Laudon & Laudon, 2010). Wijsheid vereist kritisch vermogen van de persoon die de informatie gebruikt (Rowley, 2007).

De ontwikkeling van datamanagement heeft onder meer geleid tot de volgende methoden en technieken:

- *Datawarehousing* is ontwikkeld in de jaren tachtig en

maakt het mogelijk verschillende technische platformen en datastructuren te integreren (Wixom & Watson, 2001). De data in een datawarehouse komt, net als big data, voort uit verschillende processen in de organisatie. Deze data wordt vervolgens opgeslagen in een of meerdere informatiesystemen die gerelateerd zijn aan een datawarehouse. Big data is een verlengde van de data die voortkomt uit de processen in de organisatie, waardoor we veronderstellen dat kritische succesfactoren voor de implementatie van een datawarehouse ook gelden voor de implementatie van een big data-strategie.

- *Business intelligence* (BI) is de transformatie van ruwe data naar waardevolle informatie door middel van patroonherkenning in historische dataverzamelingen (Chen et al., 2012; Laudon & Laudon, 2010; Shmueli & Koppius, 2011). Big data analyse wordt als een uitbreiding van business intelligence analyse gezien (Chen et al., 2012), waardoor we veronderstellen dat de kritische succesfactoren zoals beschreven in de literatuur met betrekking tot BI-onderzoek (Yeoh & Koronios, 2010; Işik et al., 2013) ook gelden voor big data analyse.
- *Voorspellende analyse* is ontdekken van mogelijke gebeurtenissen door analyse op historische en huidige (realtime) dataverzamelingen (Shmueli & Koppius, 2011).

2.1.2 Karakteristieken van Big Data

Er is geen unieke definitie van ‘big data’, maar een vijftal aspecten die betrekking hebben op big data worden in de literatuur genoemd: volume, snelheid, variëteit, betrouwbaarheid en waarde. Deze karakteristieken leiden tot toenemende complexiteit, omdat organisaties te maken hebben met veel data die snel op hen af komt in verschillende, complexe formaten vanuit verschillende bronnen (Franks, 2012).

Gantz & Reinsel (2011) beweren dat big data niet nieuw is, maar dat het een dynamisch concept is, mogelijk gemaakt door technische innovaties, namelijk de digitalisering van de samenleving en innovatieve geavanceerde analysetechnologieën en -technieken die het voor organisaties mogelijk maken inzichten te verkrijgen uit data in termen van snelheid en nauwkeurigheid, wat voordien zonder die technieken niet mogelijk was. Deze inzichten kunnen alleen worden verkregen als data op een adequate manier wordt verzameld en opgeslagen binnen de organisatie.

Volume

Iedere dag wordt een extreem grote hoeveelheid data gecreëerd die exponentieel toeneemt. Individuele personen creëren iedere dag data door te communiceren, kopen, delen, browsen en zoeken (Manyika et al., 2011). In een rapport van New Vantage Partners (2012) wordt

de groei van data verklaard uit de explosie van multimedia content, zoals social media, video's en foto's, in aanvulling op traditionele apparaten zoals computers.

Snelheid

Volgens onderzoeksorganisatie Gartner, omvat datasnelheid twee aspecten: de snelheid om data te verzamelen en vast te leggen en de snelheid om vervolgens de data te verwerken om deze te kunnen gebruiken voor het nemen van business beslissingen (Petthey & Goasduff, 2011). Snelheid is met name van belang, omdat tijdens het verwerkingsproces al weer nieuwe data beschikbaar komt die mogelijk andere beslissingen tot gevolg hebben (Davenport et al., 2012).

Variëteit

Datavariëteit kan in twee hoofdgroepen worden onderverdeeld: gestructureerde en ongestructureerde data. Relatief nieuwe databronnen zijn vaak ongestructureerd en vormen een belangrijk deel van big data. Gestructureerde, relationele data is georganiseerd in rijen en kolommen en het is duidelijk waar deze data vandaan komt en hoe deze kan worden gerelateerd aan acties en te nemen beslissingen. Voor de opslag en analyse van deze data wordt veelal datawarehouse-technologie gebruikt. Ongestructureerde data is vaak gegenereerd via het Internet: op sociale netwerken geplaatste berichten en updates, GPS-signalen, e-mails, multimedia content zoals video's, foto's en audio's, aflezing uit sensoren, etc. (Davenport et al., 2012). Traditionele databases c.q. datawarehouses ontberen vaak de mogelijkheid voor het uitvoeren van big data-analyse. De uitdaging voor organisaties is het combineren van traditionele gestructureerde en evolutionaire ongestructureerde data (Zikopoulos et al., 2012). Een organisatie moet beschikken over de juiste informatiesystemen die de verscheidenheid aan data kunnen opslaan en analyse tools die de verschillende data met elkaar kunnen combineren om te kunnen analyseren.

Betrouwbaarheid

Big data komt gedeeltelijk voort uit bronnen die buiten de invloed van de organisatie liggen. Hierdoor dient de organisatie aandacht te schenken aan de correctheid en accuraatheid van deze data. Betrouwbaarheid betekent dat de data overeenkomstig de werkelijkheid moet zijn (Sathi, 2012). Specifieke data van het internet, van bijvoorbeeld sociale netwerken, kan onjuist worden geïnterpreteerd en heeft daarom context nodig alvorens deze kan worden getransformeerd in nuttige en juiste informatie voor de organisatie.

Waarde

Big data bevat potentiële voordelen voor een organisatie. Een onderzoek van het McKinsey Global Institute (Manyika et al., 2011) laat zien dat data waarde kan creëren in termen van verhoogde productiviteit en con-

currentiekracht voor bedrijven en dat dit tevens economisch voordeel oplevert voor consumenten. Zij schatten dat een retailer zijn operationele marge kan verhogen met meer dan 60% als big data volledig wordt benut (zie ook paragraaf 2.2.2.).

2.2. Retail

2.2.1 Karakteristieken van retailorganisaties

Kenmerkend voor retailorganisaties is dat zij producten verkopen aan individuele klanten. Het gedrag en de voorkeuren van deze klanten zijn hierdoor één van de belangrijkste informatiebehoeften van retailers. Marketingonderzoek en datamanagement, gericht op het begrijpen van klantgedrag, zijn kritische activiteiten binnen een retailorganisatie. Deze activiteiten zijn traditioneel gezien gericht op het beheren van onder andere assortiment en voorraad (Sorescu et al., 2011).

De digitale en technologische ontwikkelingen hebben twee consequenties. Aan de ene kant hebben consumenten een betere mogelijkheid om zichzelf te oriënteren en informeren en aan de andere kant wordt door deze oriëntatiefase meer data gecreëerd over de specifieke consument en zijn voorkeuren. Dit heeft tot gevolg dat het lastiger is voor een retailer om een consument voor zich te winnen, omdat de consument veel informatie en vergelijkingsmateriaal heeft over de verschillende producten en aanbieders die op de markt actief zijn. Aan de andere kant is er door deze oriëntatiefase meer bekend over de desbetreffende consument. Retailers kunnen hier voordeel uit halen door deze consument-gegenereerde data om te zetten in slimme business modellen en gerelateerde acties.

2.2.2 Kansen in retail

In een studie van het McKinsey Global Institute identificeert Manyika et al. (2011) een aantal mogelijkheden op het gebied van marketing, merchandising, operatie, supply chain en business modellen die retailorganisaties kunnen benutten door het gebruik van big data. In tabel 1 zijn de mogelijkheden voor retailers door gebruik te maken van big data, gegroepeerd in vijf categorieën, weergegeven.

Drie toepassingsmogelijkheden worden door de Nederlandse GfK (een organisatie voor marktonderzoek) genoemd als hoofdonderwerpen voor de retail waarin big data een grote rol kan spelen (Wolters et al., 2013). Door gebruik te maken van deze mogelijkheden kan een retailer zijn bedrijfsprestaties verhogen. De eerste mogelijkheid is cross-channel sales, wat inhoudt dat verschillende verkoopkanalen aan elkaar worden gelinkt en elkaar aanvullen (bijvoorbeeld door het bestellen van een product online en afhalen of terugbrengen in de winkel). De tweede toepassingsmogelijkheid is het verbeteren van de

consumentrelevantie. Dit houdt in dat het juiste product, op het juiste moment en op de juiste plaats aanwezig is. Daarnaast kan micro-segmentatie worden toegepast, wat een meer individuele aanpak richting consumenten betreft. Mehra (2013) voegt hieraan toe dat fraudedetectie door middel van patroonherkenning in grote dataverzamelingen ook een toepassingsmogelijkheid is van big data. Bij deze toepassingsmogelijkheden worden alle beschikbare (big) data over een consument gebruikt, zoals koophistorie, real-time locatie en voorkeuren (Manyika et al., 2011). Daarnaast kan (big) data, zoals clickstream data van het web en marktonderzoek data (real-time) informatie geven over de wensen van de individuele consument.

Tabel 1 Toepassingsmogelijkheden van big data (Bron: McKinsey Global Institute, Manyika et al., 2011)

Category	Big Data opportunities
Marketing	cross selling
	location based marketing
	in-store behavior analysis
	customer micro-segmentation
	sentiment analysis
Merchandising	enhancing the multichannel consumer experience
	assortment optimization
	pricing optimization
Operations	placement and design optimization
	performance transparency
	labor inputs optimization
Supply chain	inventory management
	distribution and logistics optimization
	informing supplier negotiations
New business models	price comparison services web-based markets

3 Onderzoeksofzet

3.1 Case study methode

De methode van dit onderzoek is een case study die bestaat uit een single-case analyse aangevuld met een cross-case analyse (Staehr et al., 2012). Op basis van deze case study is een conceptueel framework ontworpen dat het formuleren en toetsen van een theorie mogelijk maakt (Eisenhardt, 1989).

3.2 Datacollectie & Caseselectie

De data zijn verzameld aan de hand van meerdere bronnen die complementair zijn aan elkaar om de betrouwbaarheid te vergroten (Yin, 2009). Ten eerste is een on-

Tabel 2 Overzicht organisaties die hebben deelgenomen aan de case study

Organisatie	Aantal geïnterviewd	Functie geïnterviewde	Branche	Status	Informatie Volwassenheidsniveau
A	2	CFO/COO; Business Analyst	Boeken	Onderdeel concern	2-3
B	2	Manager Finance & Accounting; Manager Business development	Voeding	Zelfstandig, niet beursgenoteerd	2-3
C	2	Senior Vice President Group Control; Senior Manager Consumer Research	Voeding	Beursgenoteerd	4
D	1	Head of Control & Analytics	Cosmetica	zelfstandig, niet beursgenoteerd	2
E	1	Group Controller	Kleding	Onderdeel concern	1-2
F	1	CFO	Kleding	Onderdeel concern	4-5
G	2	Director of retail; Business consultant	Kleding	Onderdeel concern	4

line vragenlijst afgenomen bij de functionarissen die hebben deelgenomen aan de interviews (zie tabel 2). De online vragenlijst bestaat uit 33 meerkeuze vragen die ingaan op de status van het huidige informatiemanagement en big data in de organisaties die deelnamen aan deze studie. Vervolgens zijn face-to-face interviews afgenomen, bestaande uit open vragen die op een semi-structureerde manier zijn gesteld. De interviewvragen zijn gebaseerd op kritische succesfactoren zoals genoemd in bestaande literatuur over informatiemanagement en -methoden. Deze vooraf opgestelde lijst van kritische succesfactoren leidde het onderzoek en ondersteunde het verzamelen en analyseren van data (Yin, 2009; Staehr et al., 2012). Uit de analyse van de data is gebleken of deze vooraf opgestelde lijst met kritische succesfactoren compleet was, of factoren moesten worden aangepast, of er irrelevante factoren waren en of er factoren ontbraken.

De selectie van cases heeft plaatsgevonden in de retailsector. Retailers kunnen inspelen op de veranderende markt door gebruik te maken van big data. Retailers hebben te maken met toenemende concurrentie door online winkels en door de steeds veeleisendere en beter geïnformeerde consumenten. Hierdoor zijn retailorganisaties een relevante doelgroep om te onderzoeken. Zeven Nederlandse retailorganisaties hebben deelgenomen aan het onderzoek. De selectie heeft plaatsgevonden op basis van persoonlijke interesse en betrokkenheid van de respondenten. Deze respondenten bestaan uit elf functionarissen, waarvan drie CFO's, vier fi-

nancieel managers/controllers en vier business (data) analisten. Het vereiste voor de respondenten was dat zij kennis hebben van informatiemanagementprocessen in de organisatie en de hieraan gerelateerde financiële aspecten.

3.3 Data-analyse

Het analyseren van de data bestond uit twee delen. Ten eerste werden de cases afzonderlijk geanalyseerd, waarbij case-specifieke patronen naar voren kwamen. Daarna zijn de cases met elkaar vergeleken, zodat de overeenkomsten en verschillen tussen de cases werden ontdekt³.

4 Resultaten

De respondenten in deze studie hebben aangegeven welke issues in hun organisatie spelen met betrekking tot (big)⁴ data. Uit de interviews blijken drie soorten issues, te weten: issues m.b.t. de waarde van (big) data, issues m.b.t. de implementatie van (big) data in de organisatie en issues gericht op de klant. In tabel 3 is een overzicht per case weergegeven.

Tabel 3 Issues per case

Case	Big data issues
A	<ul style="list-style-type: none"> - Niet alle data wordt bewust verzameld, opgeslagen en beschikbaar gesteld voor de organisatie. - Big data wordt gezien als waardevol, maar kennis ontbreekt inzake transformatie in waardevolle informatie. - Capaciteit systemen en software onvoldoende voor analyseren big data. - Veel verschillende en wisselende klanten en korte klantcontacten bemoeilijken focus op individuele klant.
B	<ul style="list-style-type: none"> - Niet alle data wordt betrouwbaar, relevant en waardevol geacht. - Niet alle relevante data wordt verzameld en opgeslagen in informatiesystemen van de organisatie.
C	<ul style="list-style-type: none"> - Onvoldoende kennis inzake transformatie grote hoeveelheden data naar waardevolle inzichten, alsmede omzetting in acties. - Beschikbare data ligt verspreid over verschillende afdelingen. - Combinatie van data uit verschillende bronnen vormt uitdaging. - Privacy consument is beperkende factor voor gebruik big data.
D	<ul style="list-style-type: none"> - Intern beschikbare data ligt verspreid in de organisatie, maakt integratie met externe data onmogelijk. - Managementinformatiesysteem niet optimaal voor big data analyse. - Onvoldoende kennis welke data relevant en juist is en waar deze zich bevindt, alsmede hoe deze data zou moeten worden omgezet in relevante informatie.
E	<ul style="list-style-type: none"> - Geen heldere visie of en hoe big data waarde kan creëren voor de organisatie. - Moeilijk inzichten uit big data om te zetten in acties, zodat deze worden uitgevoerd door de juiste personen. - Doelgroep is zeer gevarieerd, maakt focus op specifieke doelgroepen en persoonlijke communicatie moeilijk. - Productassortiment verandert snel, bemoeilijkt big data-gerelateerde activiteiten zoals optimalisatie van assortiment en locatie.
F	<ul style="list-style-type: none"> - Vertalen grote hoeveelheden data in waardevolle informatie en acties en communicatie naar organisatie vormt uitdaging. - Tweede uitdaging vormt het combineren van data en data-analyses uit de hele organisatie.
G	<ul style="list-style-type: none"> - Geen helder beeld of en hoe big data waardevol kan zijn voor de organisatie. - Uitdaging om medewerkers in dagelijkse werkzaamheden activiteiten te laten uitvoeren, gebaseerd op big data.

4.1 Kritische succesfactoren

Kritische succesfactoren zijn karakteristieke, condities en variabelen die, als zij op de juiste wijze worden ondersteund, onderhouden en worden gemanaged, een significante impact hebben op het succes van de implementatie van een strategie (Leidecker & Bruno, 1984). Op basis van een literatuurstudie met betrekking tot big data, informatiesystemen en informatiemanagement zijn zestien kritische succesfactoren geïnventariseerd, waarvan wordt verwacht dat deze relevant zijn voor een succesvolle implementatie van een big data-strategie. Een succesvolle implementatie van een big data-strategie zou moeten leiden tot het succesvol gebruik kunnen maken van big data.

Per factor is op basis van de case study geïnventariseerd in hoeverre de respondenten deze factor kritisch achten voor een big data-implementatie in de eigen organisatie. Daarnaast is in de interviews nog een viertal additionele succesfactoren genoemd.

Voor de scores per factor verwijzen wij naar Nillesen (juli 2013).

1. Volwassenheid

Het volwassenheidsmodel van Gartner is gerelateerd aan het niveau van business intelligence in de onderneming en omvat mensen, vaardigheden, processen en technologie (Hostmann & Hagerty, 2010). Het volwassenheidsmodel kent vijf niveau's van volwassenheid van informatiemanagement (1 = beginniveau, 5 = zeer volwassen). Dit volwassenheidsniveau is van belang om te bepalen of de organisatie de juiste mate van volwassenheid heeft bereikt om big data op adequate wijze te kunnen gebruiken in de organisatie. In het beginniveau (niveau 1) heeft de organisatie geen duidelijke strategie voor informatiemanagement en wordt informatie aangeleverd aan individuen op het moment dat zij hierom vragen. Het hoogste niveau van volwassenheid (niveau 5) betekent dat de organisatie in hoge mate gebruikt maakt van business intelligence, dat data is gestandaardiseerd en dat het informatiebeleid en infrastructuur goed op orde is. Het niveau van volwassenheid bepaalt in welke mate een organisatie gebruik kan maken van big data. Hoe hoger het volwassenheidsniveau, hoe beter een organisatie in staat is om diverse big data te betrekken bij de bedrijfsvoering. Het volwassenheidsniveau verschilt per onderzochte organisatie (zie tabel 2).

Op basis van de resultaten van de case study kan worden geconstateerd dat 'volwassenheid' geen kritische succesfactor is voor de implementatie, maar dat de volwassenheid het beginkader aangeeft voor de mogelijkheid om big data te gaan gebruiken.

2. Eigenschappen markt

De volatiliteit van de markt en de snelheid van veranderingen in de markt zijn indicatief voor de noodzaak een big data-initiatief te starten. Het blijkt uit de interviews dat deze factor daarom niet wordt gezien als een kritische succesfactor, maar wel als een initiator voor het gebruik van big data.

3. Bedrijfscultuur

Gebruikersacceptatie, teamchampionship en ondersteuning zijn aspecten van bedrijfscultuur (McAfee & Brynjolfsson, 2012; Wixom & Watson, 2001; Yeoh & Koronis, 2010). Over bedrijfscultuur wordt in het kader van implementatie van big data verschillend gedacht. Vier respondenten achten bedrijfscultuur als een van de belangrijkste aspecten om een big data-strategie te laten slagen.

4. Management-ondersteuning

Mensen willen participeren in een nieuw initiatief, mits daarvoor management-ondersteuning in de vorm van voldoende menskracht en budget aanwezig is (Wixom & Watson, 2001; Yeoh & Koronis, 2010). Dit wordt door acht respondenten als de meest belangrijke factor aangemerkt voor een big data-strategie, omdat de daarvoor benodigde verandering in de organisatie enkel tot stand komt als het topmanagement ervan overtuigd is dat het big data-initiatief waarde zal toevoegen.

5. Visie en afstemming met bedrijfsstrategie

De langetermijnvisie van de informatiestrategie dient in lijn te zijn met de strategische bedrijfsvisie om de organisatiedoelen te bereiken. Er dient ook een duidelijk organisatiedoel te zijn geformuleerd voor een informatiemanagement-initiatief (Yeoh & Koronis, 2010).

De resultaten van de case study tonen diverse uitkomsten:

- big data kan één van de 'informele' initiatieven zijn die niet in een overall bedrijfsstrategie is opgenomen;
- omdat de gevolgen van een big data-implementatie zo groot zijn voor de organisatie, is het beter eerst een bedrijfsbrede big data-strategie te formuleren;
- big data is gewoon een logische evolutie van informatiemanagement en er dient daarom geen specifieke strategie voor te worden opgesteld.

6. Business case

De case study toont verschillende uitkomsten over de invloed van de kwaliteit van de business case op het succes van de implementatie van een big data-strategie. Uit de interviews blijkt dat het uitmaakt wie een concrete vraag stelt waarvoor analyse van big data relevant is: topmanagement of lagere niveau's. Ook is het voorstelbaar dat een big data-analyse wordt uitgevoerd zonder dat hier een business-vraag achter ligt en dat

'gewoon gekeken wordt wat er uit komt', hetgeen een aanzienlijk risico van irrelevante uitkomsten geeft.

7. Planning

Er dient een flexibele planning te bestaan waarin ruimte is voor aanpassingen, omdat de behoefte van de organisatie gedurende een implementatie kan wijzigen (Yeoh & Koronis, 2010). Vier van de zeven organisaties achten planning kritisch voor een succesvolle implementatie van een big data-strategie. Gegeven de complexiteit dient met relatief kleine veranderingen te worden begonnen die geleidelijk worden uitgebouwd. Twee organisaties merken op dat big data deel is van de leercurve van het informatiemanagement en als zodanig geen plan of 'roadmap' kent.

8. Datakwaliteit

Bedrijfsbrede datadefinities en richtlijnen verhogen de data-kwaliteit en maken data beter interpreteerbaar en begrijpelijk (Wixom & Watson, 2001; Işik et al., 2013). Deze factor wordt door zes van de zeven onderzochte organisaties kritisch tot zeer kritisch geacht: alleen met duidelijke gestandaardiseerde data-definities kan éénduidige informatie worden gegeneerd. Dit is extra belangrijk bij big data gegeven de herkomst uit verschillende interne en externe databronnen. Echter hier ligt twijfel of dit wel gerealiseerd wordt. Met name van de accuraatheid van de data en in iets mindere mate de bruikbaarheid van de data verwachten de organisaties nog onvoldoende. Dit heeft er onder andere mee te maken dat de personen die online actief zijn en data genereren niet altijd representatief zijn voor de doelgroep van een organisatie.

9. Kwaliteit van databronnen

Gebruikers zijn meer tevreden over de kwaliteit van interne databronnen dan over die van externe databronnen (Işik et al., 2011). De kwaliteit van de databronnen en de mogelijkheden om data uit verschillende bronnen te integreren bepalen de performance van de mensen die deze data gebruiken. De kwaliteit van de databronnen wordt door zes van de zeven organisaties kritisch geacht. Drie respondenten geven aan dat de data uit eigen verkoopbestanden het meest betrouwbaar zijn. Externe marktdata wordt als minder betrouwbaar ervaren, omdat deze uit steekproeven of panels afkomstig zijn, waarin niet alle groepen even goed vertegenwoordigd zijn. Social media-data en clickstream-data van websites worden het meest genoemd als nieuwe databronnen. Daarbij wordt opgemerkt dat social media-data met name voor kwalitatief onderzoek bruikbaar wordt geacht, zoals voor indicatie trends en hypothesevorming. Voor statistische analyse worden deze data als onvoldoende gezien. Het combineren van databronnen wordt als een kans gezien, maar ook als een risico, gegeven de verschillende kwaliteitsniveau's.

10. Data-eigenaarschap

Voor het nemen van juiste beslissingen is het nodig dat de bevoegdheden voor het nemen van beslissingen en de daarvoor benodigde informatie bij dezelfde persoon beschikbaar zijn (McAfee & Brynjolfsson, 2012). Data governance en data-eigenaarschap wordt door alle organisaties als kritisch beschouwd voor een big data-implementatie. Dit hangt nauw samen met de hiervoor genoemde toename van databronnen met wisselende kwaliteit. Duidelijk en eenduidig dient te zijn geregeld wie er beslist over welke data kunnen worden gebruikt voor besluitvorming in de organisatie.

11. Resources

Innovatieve informatie-initiatieven zijn dikwijls tijdsintensief en kostbaar. Het financiële budget en het aantal bij de transformatie betrokken medewerkers zijn van invloed op een business intelligence-project (Yeoh & Koronis, 2010). De helft van de respondenten acht de huidige investeringen in systemen en technologie onvoldoende. Het aantal medewerkers dat is betrokken bij een big data-initiatief alsmede de tijd die aan big data wordt besteed wordt door bijna alle organisaties onvoldoende geacht.

12. Gebruikersparticipatie

Gebruikers van informatie kunnen waardevolle input leveren aan IT-experts zodat te bouwen modellen de juiste data-eigenschappen, business rules en datadefinities bevatten (Yeoh & Koronis, 2010). De resultaten laten zien dat de respondenten deze succesfactor gemiddeld kritisch achten.

13. Samenstelling en kwaliteit van het team

Het is van belang dat IT-medewerkers bekend zijn met de businessvereisten en dat zij mensen in de organisatie helpen met het formuleren van eisen zodanig dat deze kunnen worden gerealiseerd (McAfee & Brynjolfsson, 2012). Uit de case study blijkt dat de vaardigheden van de gebruikers kritisch worden geacht, maar dat de samenstelling van een big data-analyseteam niet in zeer hoge mate als kritisch wordt beschouwd.

14. Flexibiliteit technische infrastructuur

De mate waarin de technische infrastructuur kan worden aangepast aan wijzigende gebruikersbehoeften is van belang (Wixom & Watson, 2001). Big data wordt gekenmerkt door onder meer verschillende databronnen die kunnen wijzigen, waarbij deze wijzigingen snel moeten kunnen worden ingepast in de technische infrastructuur. Uit de case study blijkt een gemiddelde score voor de noodzaak van een flexibel technisch framework. Deze factor wordt nauwelijks genoemd in de interviews, wat er op kan wijzen dat de respondenten dit niet als een kritische factor beschouwen, dan wel dat de respondenten zich niet erg bewust zijn van deze factor.

15. Systeemkwaliteit

Deze factor, die de performance van het dataverwerkend systeem zelf betreft (DeLone & McLean, 2003), is door twee organisaties expliciet genoemd als belangrijke factor bij het gebruiken van big data. Systeemkwaliteit is afhankelijk van de input van de data en het wordt daarom belangrijk geacht om de data van hoogwaardige kwaliteit in het systeem op te slaan en te verwerken.

16. Gebruik van de output

Twee items bepalen het gebruik van informatietechnologie, te weten bruikbaarheid (gebruiker heeft perceptie dat gebruik van het systeem zijn productiviteit verhoogt) en gebruiksgemak (gebruiker heeft perceptie dat gebruik van het systeem geen extra moeite kost) (Davis, 1989). Dit aspect is door twee organisaties expliciet genoemd als kritische succesfactor.

4.2 Specifieke big data kritische succesfactoren

In de interviews is een aantal succesfactoren genoemd die niet zijn geïdentificeerd aan de hand van de literatuur, maar wel relevant lijken in het kader van een big data-strategie.

17. Geavanceerde data-analyse-expertise

Vijf organisaties huren externe organisaties in voor het uitvoeren van geavanceerde data-analyse-activiteiten, zoals het analyseren van en rapporteren over zeer grote datavolumes en analyse van social media of e-commerce. Eén respondent geeft aan dat dit geen core-activiteiten zijn voor de organisatie en dat daarvoor geen interne investeringen worden gedaan.

18. Ontwerp organisatie

Drie organisaties geven in het interview aan waar in de processen en organisatie wijzigingen nodig zijn om vervolg te geven aan de resultaten uit big data-analyse. Deze wijzigingen zijn op het gebied van:

- customer service en de fysieke winkels: snellere, 24/7 en meer individuele benadering van de klant;
- logistieke processen als gevolg van cross-channel verkoop (bijvoorbeeld online aankoop en ophalen product op een fysieke locatie).

19. Coördinatie & communicatie

Door zes van de elf respondenten worden communicatie en gecoördineerde acties binnen de organisatie als zeer kritisch genoemd voor de realisatie van een nieuw initiatief zoals implementatie van een big data-strategie, zodat alle medewerkers op de hoogte zijn van de veranderingen die big data met zich meebrengt.

20. Datawaarde

Bij een aantal organisaties is in de huidige situatie al veel data aanwezig die niet wordt benut. Hoewel uit de online vragenlijst blijkt dat met name onvoldoende ca-

paciteit en vaardigheden binnen de organisatie bestaan voor data-analyse, worden in de interviews aanvullende redenen genoemd:

- onvoorspelbaarheid van klanten;
- combineren van data wordt moeilijker naarmate er meer databronnen zijn;
- personen die een product kopen zijn niet altijd dezelfde als de personen die het product gebruiken, hetgeen het risico inhoudt dat het proces van oriëntatie tot koop niet goed wordt weergegeven.

De uitkomsten tonen dat de waarde van big data voor alle organisaties relevant is. Alle respondenten geven aan dat data relevant en betrouwbaar moet zijn om hieruit waardevolle inzichten te halen. Zolang het niet duidelijk is wat de organisatie precies wil met big data, zal er geen initiatief tot implementatie worden genomen.

5 Big Data Kritische Succesfactor Framework

Het onderzoek van de zeven cases bevestigt dat er achttien kritische succesfactoren zijn, waarvan er veertien ook gelden voor andere informatiemanagementmethoden en -systemen. Vier van deze factoren kunnen als specifiek kritisch worden geacht voor het succes van een big data-strategie (zie figuur 1).

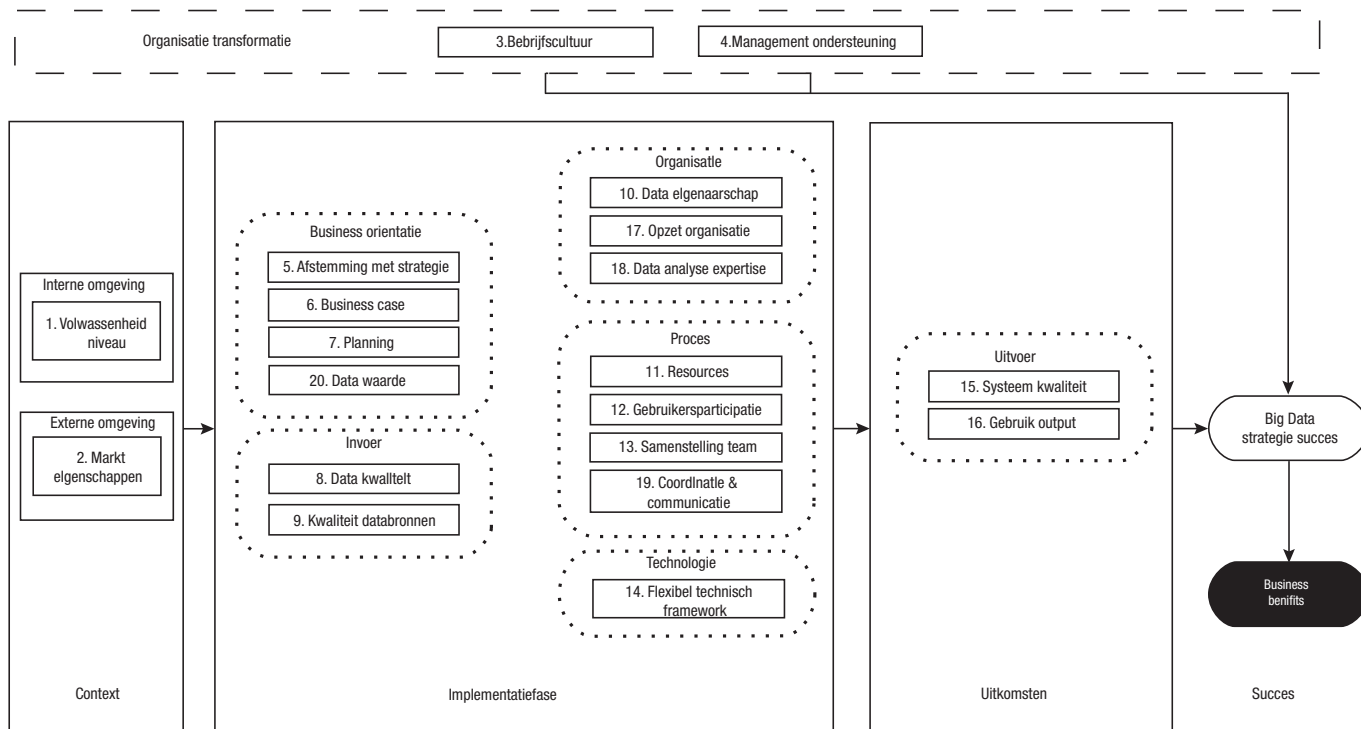
Uit de vragenlijst en interviews met de zeven retailorganisaties zijn de zestien uit de literatuur geïnventariseerde succesfactoren voor big data-strategie-implementatie bevestigd. Daarvan zijn twee factoren aangemerkt als contextueel: deze hebben geen invloed op de implementatie zelf, maar wel óf er een initiatief voor big data-implementatie zal plaatsvinden⁵. Daarnaast is door de organisaties een viertal factoren genoemd die van extra kritisch belang zijn voor een succesvolle big data-implementatie. Dit resulteert in een framework dat bestaat uit 18 kritische succesfactoren en 2 contextuele factoren. In figuur 1 is dit framework weergegeven.

De voordelen van big data voor een organisatie ('business benefits') worden gerealiseerd na een succesvolle implementatie van een big data strategie, bestaande uit de volgende vier onderliggende elementen waaronder de factoren zijn gegroepeerd⁶:

- koppeling met bedrijfsstrategie;
- invoer;
- verwerking (organisatie, proces en technologie); en
- uitvoer.

De twee contextuele factoren zijn in het grijs aangegeven. De factoren 'bedrijfscultuur' en 'managementondersteuning' beïnvloeden het hele proces van implementatie en uitvoering en zijn daarom bovenaan in het framework weergegeven.

Figuur 1 Big Data Critical Success Factor Framework



6 Conclusie

Big data is een evolutionaire stap in informatiemanagement en data-analyse en maakt gebruik van bestaande informatie-infrastructuur en informatiesystemen in de organisatie. De zestien uit de literatuur geïnventariseerde kritische succesfactoren voor informatiemanagementmethoden worden ook herkend in het kader van big data. Daarnaast zijn er vier succesfactoren onderkend die meer specifiek voor big data lijken te gelden. Deze studie geeft twee redenen voor retailorganisaties om big data te gebruiken:

- er is reeds een grote hoeveelheid data aanwezig in de organisatie die continu toeneemt door technische ontwikkelingen en de digitalisering van de samenleving, waar nieuwe inzichten uit verkregen kunnen worden;
- er zijn potentiële verbeteringen met betrekking tot de extern georiënteerde activiteiten, met name naar klanten toe, op basis van big data.

Hoewel de mogelijkheden van big data voor retailorganisaties groot zijn, wordt big data-analyse maar zeer beperkt door de onderzochte organisaties uitgevoerd. De belangrijkste redenen hiervoor zijn:

- gebrek aan managementondersteuning, vanwege onvoldoende overtuiging van de toegevoegde waarde van big data;
- het gebrek aan mogelijkheden voor het combineren van verschillende data bronnen en systemen en het analyseren hiervan.

7 Beperkingen en vervolgonderzoek

De bevindingen in dit onderzoek zijn gebaseerd op een vragenlijst en interviews met zeven organisaties, waarbij de ervaring en verwachtingen van de geïnterviewden een belangrijke rol spelen. Tevens is dit onderzoek beperkt tot retailorganisaties in Nederland. Het kritische succesfactor-framework voor de implementatie van een big data-strategie is derhalve gebaseerd op een beperkte onderzoeksgroep. Vervolgstudie dient uit te wijzen of dit framework ook van toepassing is in andere sectoren. De onderhavige studie is kwalitatief van aard en gericht op theorievorming. Door middel van kwantitatief onderzoek dienen de conclusies uit deze studie te worden getoetst. ■

Drs. J.G.I. van den Belt RA is werkzaam als universitair docent aan de post-graduate Accountancy opleiding van de Vrije Universiteit en zelfstandig consultant. C.A.C. Nillesen MSc is als consultant werkzaam bij Con-Quaestor, een onderdeel van Grant Thornton International. Dit artikel is afgeleid uit de Master thesis 'Beyond Big Data – A research in the retail sector', uitgevoerd door Charlotte Nillesen in het kader van haar Masterstudie Accounting & Control aan de Vrije Universiteit (Nillesen, juli 2013).

Noten

1 Afgeleid van een uitspraak van John Naisbitt "We have for the first time an economy based on a key resource [Information] that is not only renewable, but self-generating. Running out of it is not a problem, but drowning in it" (1988).

2 De rol van big data wordt ook wel kritisch beschouwd, getuige een uitspraak van Nassim Nicholas Taleb tijdens een door het Nexus Instituut georganiseerde masterclass. Zie: <http://www.automatiseringjds.nl/nieuws/2013/41/>

big-data-is-bullshit.

3 Voor uitgebreide informatie over de onderzoeksoepzet en aanpak verwijzen wij naar Nillesen (2013).

4 Genoemde issues hadden niet altijd specifiek betrekking op big data, daarom is 'big' tussen haakjes gezet.

5 Implementatie van een big data-strategie bij een te laag volwassenheidsniveau is niet mogelijk, omdat aan diverse organisatorische, culturele

en technische eisen niet is voldaan. In die zin kan deze factor wel als kritisch worden gezien.

Bij markteigenschappen die het gebruik van big data niet zinvol maken, is implementatie van een big data-strategie niet relevant.

6 Deze indeling is een schematische weergave die enkel inzicht beoogt te geven in de samenhang van de factoren. De samenhang tussen de factoren is in deze studie niet verder onderbouwd.

Literatuur

- Beyer, M., & Laney, D. (2012, June 21). *The importance of 'Big Data': a definition*. Geraadpleegd op Gartner: <http://www.gartner.com/id=2057415>.
- Chen, H., Chiang, R., & Storey, V. (2012). *Business Intelligence and Analytics: From Big Data to Big Impact*. *MIS Quarterly*, 36(4), 1165-1188.
- Davenport, T., Barth, P., & Bean, R. (2012). *How 'Big Data' is different*. *MIT Sloan Management Review*, 54(1), 42-46.
- Davis, F. D. (1989). *Perceived usefulness, perceived ease of use, and user acceptance of information technology*. *MIS Quarterly*, 13(3), 319-339.
- DeLone, W.H., & McLean, E.R. (2003). *The DeLone and McLean model of information systems success: A ten-year update*. *Journal of Management Information Systems*, 19(4), 9-30.
- Earl, M. J. (1989). *Management strategies for information technology*. New York: Prentice Hall.
- Eisenhardt, M. (1989). *Building theories from case study research*. *The Academy of Management Review*, 14(4), 532-550.
- Franks, B. (2012). *Taming the Big Data tidal wave. Finding opportunities in huge data streams with advanced analytics*. New Jersey: John Wiley & Sons.
- Gantz, J., & Reinsel, D. (June 2011). *Extracting value from chaos*. Sponsored by EMC Corporation.
- Gartner (2013). *Topics: Big Data*. Geraad-

- pleegd op <http://www.gartner.com/technology/topics/big-data.jsp>.
- Hambrick, D.C., & Fredrickson, J. (2005). *Are you sure you have a strategy?* *Academy of Management Executive*, 19(4), 51-62.
 - Hostmann, B., & Hagerty, J. (2010, September 17). *ITScore Overview for Business Intelligence and Performance Management*. Geraadpleegd op Gartner: <http://www.gartner.com/id=1433813>.
 - Isjik, O., Jones, M.C., & Sidorova, A. (2011). *Business intelligence (BI) success and the role of BI capabilities*. *Intelligent Systems in Accounting, Finance and Management*, 18(4), 161-176.
 - Isjik, O., Jones, M.C., & Sidorova, A. (2013). *Business Intelligence success: The roles of BI capabilities and decision environments*. *Information & Management*, 50, 13-23.
 - Laudon, K.C., & Laudon, J.P. (2010). *Management information systems*. New Jersey: Pearson Prentice Hall.
 - Leidecker, J. K. & Bruno, A. V. (1984). *Identifying and using Critical Success Factors*. *Long range planning*, 17(1), 23-32. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Geraadpleegd op McKinsey Global Institute: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.
 - McAfee, A., & Brynjolfsson, E. (2012). *Big Data: The management revolution*. *Harvard Business Review*, 90(10), 3-9.
 - Mehra, G. (2013, May 22). *Using Big Data to prevent Ecommerce fraud*. Geraadpleegd op <http://www.practicalcommerce.com/articles/4031-Using-Big-Data-to-Prevent-Ecommerce-Fraud->.
 - Minelli, M., Chambers, M., & Dhiraj, A. (2013). *Big Data and Big Analytics. Emerging business intelligence and analytic trends for today's businesses*. New Jersey: John Wiley & Sons.
 - New Vantage Partners (2012). *Big Data Executive Survey*. Geraadpleegd op http://new-vantage.com/wp-content/uploads/2012/09/nvp_exec_summary_onepage3.pdf.
 - Nillesen, C. (juli 2013). *Beyond big data – A research in the retail sector*. Master thesis, Accounting & Control, Vrije Universiteit.
 - Pettey, C., & Goasduff, L. (2011, June 27). *Gartner Says Solving "Big Data" Challenge Involves More Than Just Managing Volumes of Data*. Stamford: Gartner. Geraadpleegd op <http://www.gartner.com/it/page.jsp?id=1731916>.
 - Ragu-Nathan, B., Ragu-Nathan, T.S., Tu, Q., & Shi, Z. (2001). *Information management (IM) strategy: the construct and its measurement*. *Journal of Strategic Information Systems*, 10, 265-289.
 - Rajterič, H.I. (2010). *Overview of business intelligence maturity models*. *Management: Journal of Contemporary Management Issues*, 15(1), 47-67.
 - Rowley, J. (2007). *The wisdom hierarchy: representations of the DIKW hierarchy*. *Journal of Information Science*, 33(2), 163-180.
 - Sathi, A. (2012). *Big data analytics: Disruptive technologies for changing the game*. Boise: MC Press Online.
 - Shmueli, G., & Koppius, O. (2011). *Predictive analytics in information systems research*. *MIS Quarterly*, 35(3), 553-572.
 - Sorescu, A., Frambach, R., Singh, J., Rangaswamy, A., & Bridges, C. (2011). *Innovations in retail business models*. *Journal of Retailing*, 87, S3-S16.
 - Staehr, L., Shanks, G., & Seddon, P. (2012). *An explanatory framework for achieving business benefits from ERP systems*. *Journal of the Association for Information Systems*, 13(6), 424-465.
 - The Economist Intelligence Unit (2011). *Big Data: Harnessing a game-changing asset*. Economist Intelligence Unit Limited.
 - Wolters, W., Thuij, G., & Huijbers, K. (2013). *Top topics 2012/2013, Retail Nederland*. Geraadpleegd op: http://www.gfk.com/imperia/md/content/ps_benelux/publicaties/tt_retail_2012-2013_totaal_scherm.pdf
 - Wixom, B., & Watson, H. (2001). *An empirical investigation of the factors effecting datawarehousing success*. *MIS Quarterly*, 25(1), 17-41.
 - Yeoh, W., & Koronios, A. (2010). *Critical success factors for business intelligence systems*. *Journal of Computer Information Systems*, 50(3), 23-32.
 - Yin, R. K. (2009). *Case Study Research: Design and Methods (Fourth Edition)*. Thousand Oaks, CA: Sage publications.
 - Zikopoulos, P., Eaton, C., Roos, D. de, Deutsch, D., & Lapis, G. (2012). *Understanding big data. Analytics for enterprise class hadoop and streaming data*. USA: McGraw-Hill.