

MASTER'S THESIS

Concurrentievoordeel door data exploratie meer dan een strategie

de Vaal, E.

Award date:
2023

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 07. Feb. 2025

Open Universiteit
www.ou.nl



Concurrentievoordeel door data exploratie meer dan een strategie

Competitive advantage through data exploration more than a strategy

Opleiding:	Open Universiteit, faculteit Bètawetenschappen Masteropleiding Business Process Management & IT
Programme:	Open University of the Netherlands, faculty of Science Master of Science Business Process Management & IT
Cursus:	IM0602 Voorbereiden Afstuderen BPMIT IM9806 Afstudeeropdracht Business Process Management and IT
Student:	Edwin de Vaal
Datum:	14 mei 2023
Afstudeerbegeleider	Prof. dr.ir. Remko Helms
Meelezer	Dr. Jeroen Baijens
Derde beoordelaar	<indien aanwezig>
Versie nummer:	2.0
Status:	<definitief>

Abstract

Het hebben van een data analytics strategy of data analytics als bedrijfsdoelstelling biedt geen garantie voor het doen van data exploratie, grasduinen door de data op zoek naar innovatieve ideeën. In de praktijk gaat de prioriteit vaak naar data exploitatie, het gebruik van interne data om vragen die nu spelen te beantwoorden, bijv. verkoopcijfers, type klanten en debiteurenrisico. Dit is nuttig en nodig voor de bedrijfsvoering, maar betekent ook dat innovatieve ideeën, producten, verbeteringen in de processen niet of zeer moeilijk geïdentificeerd worden, waardoor uiteindelijk er geen concurrentievoordeel behouden kan worden of zal worden verkregen. Hoe kan een organisatie wel een verhouding tussen data exploitatie en exploratie verkrijgen?

Het hebben van een focus binnen data analytics op zowel de huidige (data exploitatie) als de toekomstige situatie (data exploratie) wordt data ambidexterity genoemd. Het vinden van een balans is in de praktijk erg lastig. Contextuele data ambidexterity wordt in de praktijk gebruikt. Afhankelijk van de omstandigheden en behoeften wordt de focus gelegd op data exploitatie of exploratie. Borging van data exploratie is beperkt, maar biedt wel vrijheid.

Op basis van het onderzoek worden er praktische handvatten aangereikt om wel degelijk stappen te maken richting een mogelijk concurrentievoordeel.

Sleutelbegrippen

Contextual, data, ambidexterity, exploration, analytics, strategy

Samenvatting

Big data heeft toegevoegde waarde voor innovaties en kan ook leiden tot concurrentievoordeel. Voor het behalen van die voordelen is het wel noodzakelijk dat er een data analytics strategy is of in ieder geval dat data analytics een onderdeel is van de bedrijfsstrategie. Op basis van die strategische doelstellingen kunnen er resources worden ingezet. Waarbij met name de resources moeten worden gezet op data exploratie, omdat daar vooral de innovatieve inzichten te vinden zijn, die uiteindelijk tot concurrentievoordeel kunnen leiden. Data exploratie is het grasduinen door interne data, meestal verrijkt met externe data, op zoek naar innovatieve ideeën.

Uit de praktijk blijkt dat ondanks het hebben van een data analytics strategy of data analytics als bedrijfsdoelstelling geen garantie biedt voor het doen van data exploratie. De prioriteit gaat in de praktijk vaak naar data exploitatie, het gebruik van interne data om vragen die nu spelen te beantwoorden. Te denken valt aan de verkoopcijfers, type klanten, debiteurenrisico's etc. Dit is natuurlijk nuttig en nodig voor de bedrijfsvoering, maar betekent ook dat innovatieve ideeën, producten, verbeteringen in de processen niet of zeer moeilijk geïdentificeerd worden, waardoor uiteindelijk er geen concurrentievoordeel behouden kan worden of zal worden verkregen. Hoe kan een organisatie wel een verhouding tussen data exploitatie en exploratie verkrijgen? Deze vraag is het onderwerp van dit exploratief onderzoek.

Het hebben van zowel een focus op de huidige situatie als op de toekomstige situatie wordt ook wel ambidexterity genoemd. En in het kader van data analytics kan gesproken worden over data ambidexterity, waaronder zowel data exploitatie als data exploratie vallen. Het vinden van een balans is in de praktijk erg lastig.

Het begrip data ambidexterity is niet onderzocht, terwijl voor organisaties dit wel van toegevoegde waarde kan zijn, zeker als men op zoek is naar concurrentievoordeel. Na het literatuuronderzoek is er middels een case study verder oriënterend onderzoek verricht.

Drie organisaties uit de financiële sector hebben meegewerkt aan deze case study en ondanks enkele verschillen waren er duidelijk veel overeenkomsten. De organisaties zijn geen directe concurrenten van elkaar en hebben hun eigen producten. Er zijn semigestructureerde interviews gehouden met twee medewerkers van elke organisatie. De medewerkers hebben alle relevante ervaring en werken langer dan een jaar bij de organisaties.

Prioriteit zal altijd liggen bij data exploitatie, maar de organisaties zijn wel bewust van de voordelen van data exploratie en maken gebruik van contextuele data ambidexterity. Dit houdt in dat inzet van resources voor data exploratie afhankelijk is van de behoeften en mogelijkheden van de organisatie. Het ontbreken van een vaste verhouding levert flexibiliteit op, maar ook minder focus op data exploratie. Deze situatie past wel bij de organisaties uit de case study.

Op basis van het onderzoek zijn er nog wel wat praktische handvatten ontdekt om wel degelijk stappen te maken richting een mogelijk concurrentievoordeel. De strategie moet een onderscheid maken tussen data exploitatie en data exploratie, zodat senior management en de organisatie het belang inzien van data exploratie, governance ten aanzien van het niet gebruiken van analyseresultaten en de inzet van artificial intelligence om het bezwaar van tijdsinvestering weg te nemen.

Summary

Big data has added value for innovations and can also lead to competitive advantage. To achieve these benefits, it is necessary that there is a data analysis strategy or at least that data analysis is part of the business strategy. Resources can be deployed based on those strategic objectives. In particular, the resources should be used for data exploration, because that is where innovative and improved solutions can be found, which can ultimately lead to competitive advantage. Data exploration is browsing through internal data, usually enriched with external data, in search of innovative ideas.

Practice shows that despite having a data analysis strategy or data analysis as a business objective does not guarantee doing data exploration. In practice, the priority often goes to data use, the use of internal data to discuss current issues. This includes sales figures, type of customers, debtor risks, etc. This is of course useful and necessary for business operations, but it also means that innovative ideas, products or process improvements are not or with a lot of effort identified. Ultimately this means that no competitive advantage can be maintained or will be obtained. How can an organization obtain a ratio between data usage and exploration? The question is the subject of this exploratory study.

Having a focus on both the current situation and the future situation is called ambidexterity. And in the context of data analytics the term 'data ambidexterity' can be used, which includes both data exploitation and data exploration. Finding a balance is very difficult in practice.

The concept of data ambidexterity has not been studied, although this can be of added value for organizations, especially if they are looking for a competitive advantage. After the literature research, further exploratory research was carried out by means of a case study.

Three organizations from the financial sector contributed to this case study and despite some differences, there were clearly many similarities. The organizations are not direct competitors of each other and have their own products. Semi-structured interviews were conducted with two employees from each organization. The employees have all relevant experience and have worked for the organizations for more than a year.

Priority will always be on data exploitation, but the organizations are aware of the benefits of data exploration and use contextual data ambidexterity. This means that the use of resources for data exploration depends on the needs and capabilities of the organization. The lack of a fixed ratio provides flexibility, but also less focus on data exploration. This situation does fit the organizations in the case study.

Based on the research, some practical tools have been discovered to take steps towards a possible competitive advantage. The strategy must distinguish between data exploitation and data exploration, so that senior management and the organization understand the importance of data exploration, governance regarding the non-use of analysis results and the use of artificial intelligence to remove the objection of time investment.

Inhoudsopgave

Abstract	ii
Sleutelbegrippen	ii
Samenvatting	iii
Inhoudsopgave	v
1. Introductie	1
1.1. Achtergrond	1
1.2. Probleemstelling	1
1.3. Onderzoeksvragen	2
1.4. Motivatie / relevantie	3
1.5. Aanpak in hoofdlijnen	3
2. Theoretisch kader	4
2.1. Onderzoeksaanpak.....	4
2.2. Uitvoering.....	5
2.2.1. Doel van het literatuuronderzoek (stap 1).....	5
2.2.2. Literatuuronderzoek (stap 3)	5
2.2.3. Practical screen (stap 4)	6
2.2.4. Quality appraisal (stap 5)	7
2.2.5. Data extraction (stap 6)	7
2.2.6. Synthese en beschrijving (stappen 7 en 8)	7
2.3. Resultaten en conclusies.....	7
2.3.1. (Big) Data analytics strategy.....	8
2.3.2. Ambidexterity	8
2.3.3. Balans bij data ambidexterity	9
2.3.4. Bevindingen.....	10
2.4. Doel van het vervolgonderzoek	10
3. Methodologie.....	11
3.1. Conceptueel ontwerp: keuze van onderzoeksmethode(n)	11
3.2. Technisch ontwerp: uitwerking van de methode	11
3.3. Gegevensanalyse.....	12
3.4. Reflectie t.a.v. validiteit, betrouwbaarheid en ethische aspecten	13
3.4.1. Validiteit en betrouwbaarheid.....	13
3.4.2. Ethiek	14

4. Resultaten	14
4.1. Uitvoering van het onderzoek	14
4.2. De resultaten.....	15
4.2.1. Eerste subvraag.....	15
4.2.2. Tweede subvraag	16
4.2.3. Derde subvraag	17
5. Discussie, conclusies en aanbevelingen	18
5.1. Discussie – reflectie.....	19
5.2. Conclusies	20
5.3. Aanbevelingen voor de praktijk.....	21
5.4. Aanbevelingen voor verder onderzoek.....	21
Referenties	23
Bijlage 1. Codes	25
Bijlage 2. Analyseresultaat.....	26

1. Introductie

1.1. Achtergrond

Big data heeft toegevoegde waarde voor innovaties (Niebel, Rasel & Viète, 2018; Duan, Cao & Edwards, 2018), maar leidt ook tot concurrentievoordeel (Lavalle, Lesser, Schockley, Hopkins, & Kruschwitz, 2011; McAfee & Brynjolfsson, 2012) en heeft een positief effect op de productiviteit (Müller, Fay & Vom Brocke, 2018).

Om die voordelen te behalen is het noodzakelijk om een heldere data analytics strategy te hebben, die een afgeleide is van de bedrijfsstrategie (Maritz, Eybers & Hattingh, 2020). Een holistische benadering is belangrijk en het moet duidelijk zijn in hoeverre de data analytics strategy het beste kan bijdragen aan de strategische doelstellingen.

Strategische doelstellingen zijn vooral op de langere termijn gericht en kunnen bijvoorbeeld gericht zijn op het behalen van concurrentievoordeel, innovaties en de productiviteit verhogen. Data analytics kan hier zeker een rol in spelen. Zeker als data exploratie wordt gebruikt, waarbij er gebruik gemaakt wordt van verrijkte data.

In de praktijk wordt data analytics vooral gebruikt voor het beantwoorden van vragen die in het hier en nu spelen (Kiron & Shockley, 2011), ook wel data exploitatie genoemd. Data exploitatie is echter minder gericht op de toekomst en kan ook niet beschouwd worden als strategische doelstelling. De beschikbare resources worden ingezet op een wijze die niet bijdraagt aan de doelstellingen uit de data analytics strategy. Dit heeft onder andere als mogelijke consequentie dat er onvoldoende concurrentievoordeel wordt behaald in de toekomst of wellicht dat het huidige concurrentievoordeel verloren gaat vanwege concurrenten die wel toekomstgerichte data analyses uitvoeren. Zeker omdat blijkt dat data toegevoegde waarde kan hebben (Lavalle, Lesser, Schockley, Hopkins, & Kruschwitz, 2011; McAfee & Brynjolfsson, 2012).

In dit onderzoek zal gekeken worden hoe een organisatie kan komen tot een optimale verdeling tussen data exploitatie en data exploratie binnen de data analytics strategy om de strategische doelstellingen te behalen.

1.2. Probleemstelling

Het kan bij organisaties voorkomen dat er een mismatch is tussen de data analytics strategy en de uitvoering daarvan. Concreet houdt het in dat er wel een datastrategie is, maar de data niet of beperkt ingezet worden om lange termijn strategische doelen te halen. Organisaties gebruiken de data vooral voor het beantwoorden van vragen over de huidige situatie, ook wel data exploitatie genoemd. Echter, wil men ook op de langere termijn overleven en succesvol zijn, dan is ook data exploratie nodig.

Voorop het gebied van data exploratie en daaruit voortvloeiend het verkrijgen of behouden van concurrentievoordeel lijkt nog wel eens uit het oog te worden verloren. Data exploratie wordt geacht tot concurrentievoordeel te leiden en wordt gekenmerkt door het verrijken van de interne data met externe data. Als een organisatie zich beschouwd als datagedreven, zijn de resultaten over

het algemeen beter op zowel financieel als operationeel vlak. Het gebruik van data exploratie lijkt te resulteren in een grotere kans op het behalen of behouden van concurrentievoordeel (Lavallo, Lesser, Schockley, Hopkins, & Kruschwitz, 2011; McAfee & Brynjolfsson, 2012).

In de huidige benaderingen is er te weinig oog voor een onderscheid tussen data exploitatie en exploratie. Het wordt meestal op één hoop gegooid onder de noemer van (big) data analysis, maar dat leidt niet automatisch tot gerichte acties om de strategische doelstellingen te halen.

De probleemstelling luidt dan ook: organisaties gebruiken data vooral nog met het oog op data exploitatie, waardoor er te weinig focus ligt op data exploratie met als gevolg dat de doelstellingen uit de data analytics strategy niet of onvolledig worden gehaald.

1.3. Onderzoeksvragen

De data analyses zijn niet altijd in lijn met de doelstellingen uit de data analytics strategy. De resources worden in de praktijk vaak ingezet op data exploitatie in plaats van data exploratie (Kiron & Shockley, 2011). Een interessante lens om dit verder te onderzoeken, alsmede om wellicht tot een betere aansluiting te komen tussen het gebruik van de data en de strategie, kan zijn ambidexterity. Ambidexterity is de zienswijze dat een organisatie naast de exploitatie ook exploratie zou moeten toepassen om langdurig concurrentievoordeel te houden. Dit kan op organisatieniveau zijn (Junni, Sarala, Taras & Tarba, 2013; O'Reilly & Tushman, 2013), maar ook op IT niveau.

Wellicht is deze theorie ook toe te passen op data, welke we hierna data ambidexterity (DA) noemen. Dit zou betekenen dat organisaties meer het onderscheid moeten maken tussen data exploratie en data exploitatie met als doel om concurrentievoordeel te krijgen of te houden.

In veel onderzoek en modellen wordt ervan uitgegaan dat big data voor zowel de exploitatie als ook voor exploratie gebruikt kan worden (Niebel, Rasel & Viète, 2019). Beide benaderingen van data worden samengevoegd en men noemt het een datastrategie. Echter, in de praktijk wordt er bij veel organisaties bewust of onbewust toch meer de focus gelegd op data exploitatie. Terwijl het niet ondenkbaar is dat in bepaalde typen organisaties, bijvoorbeeld innovatieve bedrijven, vooral de focus zou moeten liggen op data exploratie om het concurrentievoordeel te behouden of te verkrijgen.

Doordat data exploitatie en exploratie in de literatuur vaak samen onder de noemer big data worden genomen, ontbreekt de aansluiting tussen wetenschappelijk onderzoek en de praktijk. Het ontbreken van specifiek onderzoek op dit vlak is een gemis, omdat hierdoor de aansluiting tussen wetenschappelijk onderzoek en de praktijk ontbreekt.

Om de probleemstelling nader te onderzoeken zijn een aantal onderzoeksvragen geformuleerd. De probleemstelling geeft aan dat organisaties data vooral gebruiken met het oog op data exploitatie, waardoor er te weinig focus ligt op data exploratie met als gevolg dat de doelstellingen uit de data analytics strategy niet of onvolledig worden gehaald.

Dit is direct ook de belangrijkste onderzoeksvraag: hoe zorgt een organisatie voor een verhouding tussen data exploitatie en data exploratie (data ambidexterity) in de data analytics strategy?

Om tot beantwoording te komen van de onderzoeksvraag, zijn er ook een aantal subvragen geformuleerd die ondersteuning bieden bij het beantwoorden:

- Welke rol speelt data in het vaststellen van de data analytics strategy binnen een organisatie?
 - o Deze vraag probeert inzicht te krijgen in het gebruik en het soort data in de organisatie.
- In hoeverre wordt hierin onderscheid gemaakt tussen interne en externe data?
 - o Bij met name data exploratie is het gebruik van externe data belangrijk, aangezien het tot nieuwe inzichten kan leiden.
- In hoeverre is er een zeker optimum in het gebruik van interne en externe data?
 - o Als er gestreefd wordt naar een optimum tussen het gebruik van in- en externe data dan lijkt het of er sprake is van een aansluiting met de data analytics strategy.

De antwoorden op de onderzoeksvragen leiden mogelijk tot een antwoord op de vraag wat organisaties kunnen doen om hun datastrategie beter aan te laten sluiten op de gewenste strategische doelstellingen.

1.4. Motivatie / relevantie

Onderzoek toont aan dat bedrijven die zowel exploitatie als exploratie inzetten een grotere kans op concurrentievoordeel hebben, dan bedrijven die zich enkel richten op de exploitatie van de resources (Junni, Sarala, Taras & Tarba, 2013), al is dit wel afhankelijk van de context.

Het lijkt erop dat veel organisaties de focus hebben op data exploitatie (Kiron & Shockley, 2011), wat inhoudt dat men bij data analyses vooral gebruik maakt van beschikbare data en zodoende antwoorden probeert te vinden op vragen die spelen binnen de organisatie. Het ontbreekt aan focus op data exploratie en het gebruik van externe data, waardoor er minder kans is op het behalen van concurrentievoordeel door innovaties.

Er is geen onderzoek bekend over de verhouding van data exploitatie en data exploratie om concurrentievoordeel op te leveren, oftewel in lijn zijn met de data analytics strategy. Als daar meer over bekend is dan kunnen organisaties daar hun voordeel mee doen en een betere aansluiting bewerkstelligen tussen de data analytics strategy en data ambidexterity. Het voordeel is met name gelegen in het halen van de strategische doelstellingen.

Naast het halen van de strategische doelstellingen zou het ook kunnen leiden naar het doeltreffender inzetten van resources, zodat er mogelijk een efficiencyslag of kostenbesparing gemaakt kan worden. Er zou bijvoorbeeld geen externe data gekocht hoeven te worden omdat de focus ligt op data exploitatie.

1.5. Aanpak in hoofdlijnen

In sectie 1.3 zijn de onderzoeksvraag en sub vragen geformuleerd en in deze sectie zal in hoofdlijnen het verdere verloop van het onderzoek beschreven worden.

In hoofdstuk 2 wordt het theoretisch kader behandeld, met daarin het literatuuronderzoek en de uitkomsten.

Er zal exploratief onderzoek worden gedaan, aangezien we willen weten wat er gebeurt bij organisaties. Als eerste stap in het exploratief onderzoek zal er een literatuuronderzoek plaatsvinden, waarna er een aantal semigestructureerde interviews worden afgenomen.

De opzet en onderzoeksmethodiek van het literatuuronderzoek en de interviews worden in hoofdstuk 3 in meer detail weergegeven, waarna in hoofdstuk 4 alle resultaten worden behandeld. In het laatste hoofdstuk, hoofdstuk 5, zal er nader worden ingegaan op mogelijke conclusies en aanbevelingen.

2. Theoretisch kader

2.1. Onderzoeksaanpak

Het literatuuronderzoek maakt gebruik van de methodiek beschreven door Okoli en Schabram (2010). De focus van het gedetailleerde achtstappenplan ligt vooral op het gebied van information systems, dus dat sluit aan bij dit onderzoek. Een groot voordeel van het stappenplan is dat er getracht wordt om naast de praktische toepasbare stappen ook duidelijk wordt verwezen naar de theorie en de academische literatuur. Het stappenplan vergroot de transparantie, alsmede de reproduceerbaarheid, omdat alle stappen systematisch worden vastgelegd.

De eerste stap uit het stappenplan begint bij het vaststellen van het doel van het onderzoek. In het geval er meerdere onderzoekers zijn, dan is het raadzaam om een onderzoeksprotocol op te stellen. In dat protocol wordt vastgelegd, hoe de zoekcriteria worden toegepast, hoe relevante academische literatuur wordt vastgelegd etc. Dit is stap twee en wordt overgeslagen, aangezien het onderzoek wordt uitgevoerd door één onderzoeker.

De derde stap behelst het zoeken naar de relevante literatuur en het vastleggen van de resultaten. Nadat de resultaten gevonden zijn, zullen de onderzoeken beoordeeld worden op relevantie. Deze vierde stap wordt ook de 'practical screen' genoemd.

De gevonden resultaten worden niet alleen beoordeeld op praktische relevantie, maar ook op kwaliteit, de 'quality appraisal'. Dit is de vijfde stap.

In de zesde stap zal er op basis van de gevonden en beoordeelde literatuur bekeken worden welke elementen uit de verschillende onderzoeken relevant zijn. Dit wordt ook wel de data extraction fase genoemd. In deze fase wordt ook data vastgelegd over het onderzoek, zoals het jaar van publicatie, de wijze van verzameling, informatie over waarom een onderzoek juist beter niet gebruikt kan worden etc. Dit wordt allemaal vastgelegd, zodat op een later tijdstip hier gebruik van kan worden gemaakt.

In de volgende stap, de zevende stap, worden de resultaten geanalyseerd. Er wordt beoordeeld wat de samenhang is tussen de gevonden onderzoeken. De eerder genoemde vastlegging kan hierbij ondersteuning leveren. Ook is getracht om het conceptueel model te linken met de gevonden

literatuur. Dit leverde meer inzicht op wat exact de relatie en relevantie is van de gevonden literatuur voor dit onderzoek.

Als laatste en achtste stap zullen alle bevindingen beschreven worden en zal de review verder worden uitgeschreven. Dit resultaat is terug te vinden in hoofdstuk 2.3.

2.2. Uitvoering

2.2.1. Doel van het literatuuronderzoek (stap 1)

In deze stap is gekeken naar het doel van het literatuuronderzoek in combinatie met de gestelde onderzoeksvragen. Het doel van dit literatuuronderzoek is om te bepalen of er al onderzoek is gedaan naar data ambidexterity en de data analytics strategy. Ook wordt het literatuuronderzoek gebruikt voor het bepalen van definities.

Het opstellen van een onderzoeksprotocol, stap twee, wordt overgeslagen, omdat er één onderzoeker is en een onderzoeksprotocol vooral toegevoegde waarde heeft bij meerdere onderzoekers zodat de werkwijze consistent blijft.

2.2.2. Literatuuronderzoek (stap 3)

Er is vooral gezocht in de bibliotheek van de Open Universiteit. Middels deze bibliotheek is er toegang tot verschillende andere databases en bibliotheken voor wetenschappelijk onderzoek. Als er een andere bibliotheek is gebruikt dan wordt dat expliciet gemeld,

Er zijn verschillende zoekopdrachten gebruikt om te komen tot een bruikbaar literatuuronderzoek. De query's samenvoegen om tot één integrale query te komen, was lastig omdat de invalshoeken verschilden. Voorts waren er na filtering behapbare aantallen over (Tabel 1).

Tabel 1

Gebruikte initiële zoekopdrachten en de resultaten

Query	Resultaat (#)	Resultaat na filters (#)	Gebruikte resultaten voor data extraction (#)	Gebruikt voor theoretisch kader (#)	Opmerking/Reden
(data) AND (analytics) AND (strategy)	68.509	3.762	15	5	Deze zoekterm is onderdeel van de hoofdvraag. En nodig voor het begrip
ambidexterity	2.611	104	6	2	Het principe en definitie werd hiermee onderzocht
(IT) AND (ambidexterity)	2.574	104	12	3	IT ambidexterity was mogelijk een basis voor data ambidexterity
(data) AND (ambidexterity)	1.200	67	8	3	Bepalen of er onderzoek naar is gedaan. Dit is de lens die

					wordt gebruikt voor het eigen onderzoek
(ambidexterity) AND (in practice)	551	27	5	2	Met deze zoekopdracht werd bekeken hoe in de praktijk ambidexterity wordt gebruikt en of er lessen uit getrokken kunnen worden voor data ambidexterity
(ambidexterity) AND (data) AND (analytics)	57	n.v.t.	8	3	De resultaten waren beperkt, daarom geen filter toegepast.

Noot. In de kolom Resultaat staan de aantallen gevonden na het invoeren van de zoektermen. Als de initiële aantallen laag waren dan was er geen verdere filtering nodig.

In Tabel 1 wordt ook het resultaat na filtering weergegeven. Er is gebruik gemaakt van het filter op vakgebied. Het filter op het vakgebied 'computer science' wordt gebruikt. Ook wordt er enkel gezocht op Engelse teksten en op 'Peer-reviewed publicaties'.

Verder kan het voorkomen dat ondanks de filtering toch nog veel resultaten overblijven. Dan helpt het om op basis van termen onderwerp gebieden uit te sluiten. De volgende onderwerpen werden uitgesloten: 'engineering', 'engineering', 'electrical & electronic', 'chemistry, analytical', 'electrical engineering', 'life sciences & biomedicine', 'artificial intelligence' en 'chemistry'.

Naast het zoeken in de database van de OU, is er ook gebruik gemaakt van snowballing. Hiervoor werd het 'Web of science' gebruikt. Vooral de analysefunctie van Web of science levert soms nieuwe(re) en relevante artikelen op. Ook backwards snowballing op basis van de gevonden literatuur is toegepast.

De website 'Connected Papers' (<https://www.connectedpapers.com/>) is gebruikt om gerelateerde onderzoeksliteratuur te vinden. Deze site linkt een gevonden onderzoek met andere onderzoeken, waardoor er ook gerelateerde onderzoeken gevonden kunnen worden.

2.2.3. Practical screen (stap 4)

Het sorteren van de resultaten vindt plaats op relevantie, dus dan is op basis van de titel een selectie gemaakt van de te bekijken literatuur. Enkel de eerste honderd zoekresultaten zijn bekeken, omdat daarna de relevantie tot nagenoeg nul was afgenomen. Nadat er een eerste selectie is gemaakt op basis van de titel, zijn vervolgens de abstracts van de overgebleven artikelen beoordeeld op relevantie. De artikelen die op het eerste gezicht relevant lijken, worden bijgehouden in een lijst in Microsoft Excel.

De lijst in Microsoft Excel (hierna: de Excellijst) heeft de volgende velden:

- Zoektermen: om het resultaat te linken met de zoekopdracht
- Hoe gevonden?: hier werd bijgehouden of er wellicht sprake was van snowballing of een andere wijze van vinden.
- Nummer: elk gevonden relevant onderzoek kreeg een volgnummer. Dat nummer maakte het gemakkelijk om snel te verwijzen naar een onderzoek.
- Schrijver: de naam van de eerste onderzoeker

- Jaar: het jaar van publicatie
- Web of science: aantal citaten
- Titel: de titel van het onderzoek
- Relevantie 1 t/m 5: hier was ruimte om meerdere elementen te benoemen uit het onderzoek, wat mogelijk gebruikt kan worden in het eigen onderzoek. Dat kunnen ook elementen zijn om het onderzoek juist niet te gebruiken.

De velden aangaande relevantie worden verder aangevuld in stap 6.

2.2.4. Quality appraisal (stap 5)

De artikelen zijn daarna in meer detail doorgenomen om de kwaliteit te beoordelen. Er is gekeken of het uitgevoerde onderzoek voldoende is. Er is beoordeeld op methodiek, de grootte van het onderzoek en de resultaten.

Als een ander, maar wat minder zwaarwegend onderdeel, is er ook gekeken naar hoe vaak een artikel geciteerd is. Recente artikelen zijn soms niet geciteerd en dan wordt dit onderdeel niet meegenomen in de beoordeling van de kwaliteit.

2.2.5. Data extraction (stap 6)

In de data extraction fase, stap zes, worden de onderzoeken bekeken in het licht van de sectie 1.3 geformuleerde onderzoeksvragen. Er wordt bekeken in hoeverre bepaalde elementen uit de onderzoeken relevant zijn voor het theoretisch kader. Een element kan zijn een uitkomst van een onderzoek, de onderzoeksmethodiek, een definitie of gewoon een verwijzing naar andere literatuur. De relevante elementen worden toegevoegd aan de Excellijst, zodat er een overzicht ontstaat van alle relevante elementen of juist minder relevante elementen. Die elementen zullen uiteindelijk geanalyseerd worden om thema's te ontdekken. Ook is in deze fase kritisch bekeken of er sprake was van bias in de onderzoeken.

2.2.6. Synthese en beschrijving (stappen 7 en 8)

Zoals hiervoor beschreven zijn de relevante elementen uit de onderzoeken in kaart gebracht en in een overzicht geplaatst. Voor een verder overzicht en het vinden van verbanden zijn de belangrijkste termen en elementen op papier gezet in de vorm van een mindmap. Die mindmap heeft uiteindelijk geleid tot het theoretisch kader. In de volgende sectie zal daar dieper op worden ingegaan. Daar wordt stap acht verder uitgewerkt, namelijk het beschrijven van de resultaten.

2.3. Resultaten en conclusies

In deze sectie worden de bevindingen van het literatuuronderzoek gepresenteerd. Dit komt overeen met de stappen zeven en acht (Okoli, 2010). Zoals beschreven is in sectie 1.4 gaat het over de aansluiting tussen data, het gebruik van ambixterity met data en met de data analytics strategy. Deze termen zijn dan ook onderzocht in de literatuur.

2.3.1. (Big) Data analytics strategy

Data analytics wordt tegenwoordig vaak geschaard onder big data analytics (BDA) of worden samengevoegd (Harlow, 2018). In big data analytics wordt er zowel gekeken naar huidige issues als toekomstige ontwikkelingen (Tsai, Lai, Chao, & Vasilakos, 2015). In de afgelopen jaren is veel onderzoek gedaan naar big data analytics. Het kan helpen om de strategische koers te bepalen, concurrentievoordeel te verkrijgen, betere besluitvorming en betere performance etc. (Wang, Kung & Byrd, 2018; Popovič, Hackney Tassabehji & Castelli, 2016).

BDA vergt over het algemeen wel een strategische keuze (Harlow, 2018). Om gebruik te maken van BDA dienen er voldoende resources ingezet te worden, hardware en software en het senior management moet hier ook bij betrokken te worden (Maritz, Eybers & Hattingh, 2020). BDA wordt beschouwd als een voortvloeisel uit de Data Analytics Strategy, wat op haar beurt een voortvloeisel uit de bedrijfsstrategie zal moeten zijn. Op senior management niveau zal BDA gedragen moeten worden om succesvol te zijn en resultaten te genereren.

De Data Analytics Strategy moet de strategische doelen aangeven die bereikt moeten worden door middel van de inzet van BDA. Op basis van die doelen zal er gekeken moeten worden of data exploitatie of data exploratie nodig is, of mogelijk zelfs een combinatie. Overigens ontbreekt deze stap meestal en onderzoek hierover ontbreekt (Rialti, Marzi, Caputo & Mayah, 2020). Hier komt ambidexterity om de hoek kijken.

2.3.2. Ambidexterity

Met ambidexterity wordt vaak het spanningsveld tussen exploitatie en exploratie aangeduid (March, 1991). Het spanningsveld zal hierna worden toegelicht, maar eerst een korte uiteenzetting over de termen exploitatie en exploratie.

Exploitatie is gericht op de huidige situatie verbeteren, keuzes maken, efficiencyverbetering, selectie van alternatieven, executie en implementatie van genomen beslissingen. Exploratie is gericht op de toekomstige situatie. Termen die hierbij horen zijn: zoeken, variatie, risico's nemen, experimenteren, flexibiliteit, ontdekking en innovatie (March, 1991).

Toepassing van de term 'exploitatie' op data betekent dat er in de bestaande data gezocht wordt naar antwoorden op vragen die nu spelen binnen de organisatie. Voor 'exploratie' betekent het dat er in de data gezocht wordt naar mogelijk toekomstige innovaties. Het gebruik van externe data kan daarbij ondersteunen (Kumar & Singh, 2018; Idreos, Papaemmanouil & Chaudhuri, 2015).

BDA maakt gebruik van zowel data exploitatie als data exploratie (Tsai, Lai, Chao, & Vasilakos, 2015) en wordt altijd samen genomen in de literatuur. Dus als het gaat over BDA dan gaat het over zowel data exploitatie als exploratie. Vooral data exploratie kan voor organisaties voordelen bieden, zoals helpen om de strategische koers te bepalen en concurrentievoordeel te verkrijgen (Wang, Kung & Byrd, 2018).

Wel kent exploratie een aantal nadelen (March, 1991) :

- De uitkomst is onzeker.
- Het resultaat ligt verder weg in de tijd.
- Het is lastiger om de organisatie mee te krijgen in aanpassingen op de lange termijn.

Die nadelen leiden ertoe dat organisaties de focus leggen op data exploitatie (Kiron & Shockley, 2011). Dat betekent dat bij data analyses vooral gebruik gemaakt wordt van beschikbare data en men zodoende antwoorden probeert te vinden op vragen die spelen binnen de organisatie. De beschikbare resources, zoals data, personeel, hard- en software etc., worden ingezet op gebieden waar de focus ligt, waardoor er weinig resources overblijven om te werken aan data exploratie. En dit is het eerder genoemde spanningsveld. Er moet een keuze worden gemaakt tussen de lopende vraagstukken te beantwoorden of onderzoek doen naar mogelijk toekomstige mogelijkheden. Dit spanningsveld op het gebied van data, noemen we hierna Data Ambidexterity (DA).

Data Ambidexterity is een nieuwe term die staat voor het spanningsveld tussen data exploitatie en data exploratie. In de volgende sectie wordt verder ingegaan over de balans bij DA.

2.3.3. Balans bij data ambidexterity

Als er een spanningsveld is, dan zal er ook een balans gevonden moeten worden. Hier gaat het om de balans tussen data exploitatie en data exploratie. Deze balans is nog nooit onderzocht op het gebied van data analyse, maar er zijn wel gerelateerde onderzoeken gedaan op andere gebieden, zoals IT en op organisatieniveau. Op deze onderzoeken wordt hierna ingegaan.

Information Technology (IT) kent een soortgelijke situatie. IT exploitatie zorgt dat de huidige systemen goed en betrouwbaar werken, IT exploratie zorgt er voor dat de systemen ook toekomstbestendig zijn (Benitez, Llorens, & Braojos, 2018). Toekomstbestendige systemen betekent niet automatisch de beste systemen voor de huidige situatie.

IT ambidexterity wordt vaak gezien als een enkele samengestelde constructie, altijd in balans of in ieder geval zijn beide dimensies, IT exploitatie en exploratie, even sterk (Liang, Wang & Xue, 2022; Heckmann & Maedche, 2018). Als dit principe toegepast wordt op data dan zou dat betekenen dat er op zowel data exploitatie en data exploratie evenveel resources werken. Het is natuurlijk mogelijk, maar lijkt geen rekening te houden met de omstandigheden van dat moment, het is erg star.

Een andere mogelijkheid is rekening houden met de context, waarbij er een proportionele balans is tussen beide dimensies en waarbij de optimale balans afhankelijk is van de context. De context in het onderzoek van Liang is vooral de dynamiek van de omgeving. De dynamiek van de omgeving bepaald de balans tussen IT exploitatie en exploratie (Liang et al., 2022). Dit betekent bijvoorbeeld in de praktijk dat de balans ook afhankelijk is van hoeveel medewerkers er totaal beschikbaar zijn voor IT Ambidexterity.

Ook op organisatieniveau is er een spanningsveld, ook wel 'organisational ambidexterity' genoemd. Dat is de capaciteit van een organisatie om zowel in lijn te werken met de behoeften van de organisatie in de huidige tijd, maar de organisatie kan zich ook aanpassen aan een veranderende omgeving in de toekomst (Gibson & Birkinshaw, 2004). Het gebruik van organisational ambidexterity in dynamische en onzekere omstandigheden is positief gerelateerd aan een verhoogd innovatievermogen, betere financiële resultaten en hogere overlevingskansen (O'Reilly & Tushman, 2013).

Initieel was de gedachte om duale structuren te gebruiken om het spanningsveld op te lossen bij organisational ambidexterity. Dus bepaalde business units, of bepaalde groepen in de business units, focusten op de huidige behoefte en andere business units, of groepen in de business units, focusten

op de verandering van de organisatie. Men noemt dit ook wel structural ambidexterity (Duncan, 1976).

Ook op het gebied van organisational ambidexterity is men uiteindelijk uitgekomen op contextuele organisational ambidexterity. Er wordt zowel gewerkt aan de behoeften van de organisatie op dit moment en er wordt ook gekeken naar de toekomstige behoefte van de organisatie, zonder dat een er een extra afdeling voor wordt opgericht (Gibson & Birkinshaw, 2004).

Zou het idee van structural ambidexterity worden toegepast op data ambidexterity, dan zouden er twee afdelingen zijn die zich bezig houden met BDA. Eén afdeling past structureel data exploitatie toe en de andere afdeling data exploratie en zoals eerder al genoemd dit leidt tot een star model.

Het zou best kunnen zijn dat de balans bij data ambidexterity ook afhankelijk van de context is, net als bij verschillende onderzoeken op het gebied van IT ambidexterity en organisational ambidexterity. Er wordt gekeken naar de dynamiek van de omgeving, hoeveel resources zijn beschikbaar en waar ligt de behoefte van de organisatie. En op basis daarvan wordt er een keus gemaakt in de inzet van resources en dus ook de balans in DA (Raisch, Birkinshaw, Probst & Tushman, 2009).

2.3.4. Bevindingen

Uit het literatuuronderzoek komen een aantal bevindingen naar voren:

1. In BDA wordt zowel data exploitatie en exploratie samengenomen. Het wordt op één hoop geveegd.
2. Data ambidexterity is niet onderzocht en al helemaal niet in combinatie met de data analytics strategy.
3. De optimale balans in ambidexterity is lastig te bepalen. En ook in de literatuur zijn er verschillende zienswijzen te vinden.

In ieder geval kunnen de onderzoeksvragen nog niet afdoende beantwoord worden. Er zal een vervolgonderzoek moeten plaatsvinden.

2.4. Doel van het vervolgonderzoek

Zoals uit sectie 2.3.4 blijkt, worden de onderzoeksvragen niet volledig beantwoord door de literatuur. Het vervolgonderzoek heeft als doel de onderzoeksvragen wel te beantwoorden. De beantwoording van de onderzoeksvragen kan er wellicht toe leiden dat een data analytics strategy beter uit te voeren is als de juiste balans in DA wordt gehanteerd, met als resultaat dat strategische doelstellingen op een efficiëntere wijze behaald kunnen worden.

Mocht uiteindelijk blijken dat data ambidexterity geen toegevoegde waarde heeft, dan biedt het wellicht toch inzichten die gebruikt kunnen worden voor verder onderzoek.

Voor het vervolgonderzoek is het noodzakelijk te weten in welk soort omgeving de organisaties opereren, hoe data ambidexterity toegevoegde waarde heeft voor de organisatie, welke data er nodig is en welke elementen nog meer noodzakelijk zijn om een optimale balans te vinden te tussen data exploitatie en exploratie.

Met een vervolgonderzoek kan dus een bijdrage geleverd worden aan een theoretisch onderbelicht onderwerp en handvatten bieden voor veel organisaties om effectiever en efficiënter gebruik te maken van data, zodat de strategische doelstellingen behaald kunnen worden.

In het volgende hoofdstuk zal nader worden ingegaan op het ontwerp van het vervolgonderzoek.

3. Methodologie

In deze sectie wordt nader ingegaan op het verdere onderzoek en hoe het onderzoek eruit zal komen te zien.

3.1. Conceptueel ontwerp: keuze van onderzoeksmethode(n)

Het vervolgonderzoek betreft exploratief onderzoek om te verkennen wat data ambidexterity voor toegevoegde waarde kan zijn voor een data analytics strategy, aangezien het nog een onbekend terrein is in de literatuur. Er zal worden onderzocht hoe organisaties met deze onderwerpen omgaan in de praktijk. Doordat er concrete en contextuele informatie verzameld dient te worden over een specifiek onderwerp is besloten om een case study uit te voeren. Er zullen drie organisaties participeren met de case study. De reden om gebruik te maken van meerdere organisaties is om te onderzoeken of de resultaten met elkaar overeenstemmen. Dit is mede omdat een organisatie meestal maar één data analytics strategy heeft.

De selectie moet nog worden gedaan, maar er zal gezocht worden naar organisaties met de volgende kenmerken:

- een profit organisatie;
- de grootte zal variëren tussen de 400 en 5.000 medewerkers;
- de organisaties hebben een eigen data analytics teams;
- er is een data analytics strategy.

De verwachting is dat dit type organisatie mogelijk moeite heeft of de middelen niet heeft om daadwerkelijk ook een bewuste keuze voor data exploratie te maken.

Voor de case study zullen er interviews worden afgenomen en in principe zijn alle interviewvragen open vragen. In dit geval is de meest voor de hand liggende keus om gebruik te maken van semigestructureerde interviews. Dit type interview geeft een leidraad qua vragen, maar geeft ook de geïnterviewde de tijd om zaken toe te lichten. De vragen zullen betrekking hebben op de data analytics strategy, de inzet van de resources, de cultuur en of data ambidexterity toegevoegde waarde zou kunnen hebben.

3.2. Technisch ontwerp: uitwerking van de methode

Zoals in sectie 3.1 beschreven zal er gebruik worden gemaakt van case studies en semigestructureerde interviews. Om de kwaliteit te borgen en de informatie ook vanuit een andere hoek te bekijken, zullen er per organisatie twee verschillende personen worden geïnterviewd. De redenen hiervoor zijn: proberen een vollediger inzicht te krijgen en een verhoogde validiteit van de resultaten. De interviews zijn één-op-één interviews en mogelijk via Teams of in een fysieke meeting. Een fysieke meeting heeft de voorkeur omdat de non-verbale communicatie wellicht een gelegenheid geeft om nog betere vragen te stellen. Er wordt wel getracht om de interviews maximaal een uur te laten duren en ook op te nemen.

Om tot de juiste inzichten te komen zal er contact worden gezocht met mensen die in nauw contact staan met of werken op een data analyse afdeling of betrokken zijn bij de totstandkoming van de data analytics strategy. Hier valt te denken aan senior managers en directieleden. Om de organisatie beter te doorgronden is het wel de eis dat de geïnterviewden meer dan één jaar bij het bedrijf werken en het liefst drie of meer jaar ervaring op het vlak van data analytics.

Mocht er documentatie aanwezig zijn, dan wordt uiteraard verzocht om inzage. De vragen zullen van te voren met de geïnterviewden gedeeld worden, zodat er minder toelichting nodig is en de geïnterviewden de gelegenheid te geven om de vragen voor te bereiden. Ook zal de geïnterviewden gevraagd worden om een toestemmingsverklaring voor deelname aan een afstudeeronderzoek te ondertekenen.

De onderzoeksvragen bieden de basis voor de vragen in het interview. In Tabel 2 staan de vragen opgenomen die horen bij de onderzoeksvragen, alsmede de volgorde. De vragen gaan van algemeen naar specifiek.

Tabel 2

Vragen die worden gesteld bij de onderzoeksvragen

Vragen
1. In wat voor markt opereert de organisatie?
2. Hoe wordt er vorm gegeven aan de data analytics strategy?
3. Is er een data analytics strategy binnen de organisatie?
4. Op welke wijze maakt de concurrentie gebruik van data?
5. Hoe kan data de concurrentiepositie wel of niet verbeteren?
6. Waar binnen de organisatie wordt er gebruik gemaakt van data analytics?
7. Welk Boardlid is verantwoordelijk voor de data strategie?
8. Hoe ziet de data governance eruit?
9. Hoe zou de data analytics strategy eruitzien als ambidexterity zou zijn gebruikt bij het opstellen van de strategie?
10. Hoe wordt er een onderscheid gemaakt tussen data exploitatie en data exploratie?
11. Wat zou een goede verhouding zijn? En waarom?
12. Welke resources worden er gebruikt voor data analytics?
13. Als data ambidexterity zou worden toegepast, wat zou de consequentie zijn geweest voor de resources? Bijv. hard- en software

In sectie 3.3 wordt beschreven hoe de resultaten worden uitgewerkt en geanalyseerd.

3.3. Gegevensanalyse

De interviews zullen worden getranscribeerd en de antwoorden, oftewel de resultaten worden in een tabel geplaatst. Om het proces van transcriberen te versnellen zal er getracht worden om gebruik te maken van voice recognition software. De uitkomsten worden met elkaar vergeleken en indien mogelijk vertaald naar een makkelijker te gebruiken classificatie. Hiervoor wordt de thematic analysis gebruikt (Braun & Clarke, 2006).

Doordat tijdens het transcriberen er al bekendheid zal worden verkregen, zal de eerste stap in de analyse het coderen van de getranscribeerde tekst zijn. Coderen houdt in dat aan delen van de tekst

een code wordt gekoppeld. Nadat de tekst is gecodeerd kan deze in de tweede stap worden geanalyseerd door te zoeken naar overkoepelende thema's en/of verbanden.

In stap drie zal de onderliggende data een belangrijke rol spelen, aangezien er accurate conclusies getrokken zullen worden op basis van de data. De thema's zullen een naam krijgen die de inhoud van het thema begrijpelijk weergeeft. Dit zal de definitieve lijst met thema's zijn.

Stap vier zal bestaan uit het vaststellen en structureren van de thema's en het geven van een definitieve naam. In de laatste stap, stap vijf, zal de conclusies beschreven worden. In eerste instantie in een paar zinnen en in het onderzoeksverslag uitgebreider. Steeds zullen er checks plaatsvinden op de originele tekst om te zien of de codes en thema's een accurate weergave van de data zijn.

Het voordeel van deze methode is dat er op verschillende moment in het proces een check op de data plaatsvindt. Het nadeel is wel dat dit arbeidsintensief is en dus veel tijd kost. Wel lijkt deze methode te resulteren in een goede kwaliteit van de onderzoeksresultaten.

3.4. Reflectie t.a.v. validiteit, betrouwbaarheid en ethische aspecten

3.4.1. Validiteit en betrouwbaarheid

De kwaliteit van een onderzoek wordt uitgedrukt in validiteit en betrouwbaarheid. Er zijn verschillende soorten validiteit. Voor dit onderzoek worden de volgende types validiteit gebruikt:

- Constructvaliditeit wordt gebruikt om aan te tonen dat het onderzoeksinstrument daadwerkelijk het begrip of het construct heeft gemeten dat werd beoogd te meten, of mogelijk (gedeeltelijk) iets anders. Dit is de convergente validiteit. En de discriminante validiteit wordt getoetst door te controleren of constructen die niet zouden moeten samenhangen ook daadwerkelijk niet samenhangen.
- Interne validiteit: er moet correct geredeneerd zijn en de resultaten zijn valide.
- Externe validiteit: er moet beargumenteerd en beschreven worden in hoeverre de resultaten te generaliseren zijn.

De construct validiteit wordt geborgd door de correlatie te meten met de andere interviews (convergente validiteit) en door de correlatie te meten tussen de interview resultaten en de ontvangen documentatie of jaarrekening (discriminante validiteit).

De interne validiteit wordt geborgd door twee ervaren medewerkers te interviewen die betrokken zijn bij de (uitvoer van) de data analytics strategy. Beiden krijgen dezelfde vragen en het vaststellen van causaliteit tussen de onafhankelijke variabele en de afhankelijke variabele is een belangrijk onderdeel van de vragen.

Door drie organisaties in de financiële sector te gebruiken is er ook sprake van een goede externe validiteit naar in ieder geval kleine tot middelgrote financiële ondernemingen.

De betrouwbaarheid zal geborgd worden door de interviews met twee medewerkers en de ontvangen documentatie. Triangulatie is hierbij van belang.

Betrouwbaarheid heeft betrekking op de mate waarin een meting in de scriptie vrij is van fouten, op consistente wijze iets meet en daarmee herhaalbaar is. Dit onderzoek is op een verantwoorde manier opgezet en de gebruikte methodologie is valide. De gehanteerde methode van case study

onderzoek is een veelgebruikte methode om exploratief onderzoek te doen. De focus van het onderzoek is afgebakend.

3.4.2. Ethiek

Het onderzoek zal ethiek voorop stellen. Er zullen geen namen zonder toestemming worden gebruikt. Ook zal discriminatie van medewerkers of klantgroepen niet worden getolereerd. Indien gewenst zal er een Non-Disclosure Agreement (NDA) worden ondertekend, zodat altijd duidelijk is bij de geïnterviewden en de meewerkende organisaties wat er met de data gebeurt.

4. Resultaten

4.1. Uitvoering van het onderzoek

Initieel lag de focus op organisaties die exact zouden voldoen aan de gestelde criteria (sectie 3.1) en deze organisaties werden ook benaderd, maar in de praktijk was de combinatie van de gestelde eisen lastig te vinden. Kleinere organisaties hebben meestal geen data analytics strategy of eigen data analytics team. Ook in de bedrijfsstrategie komt niet altijd naar voren dat data een onderdeel van de strategische doelstellingen is. Uiteindelijk is besloten om de eisen ruimer te stellen.

De afwijkingen op de oorspronkelijk gestelde eisen betreffen:

- De grootte is gewijzigd: minimaal 8.000 medewerkers
- Het hebben van een data analytics strategy als eis is ruimer opgevat. Als in de bedrijfsstrategie data en data analyse een duidelijke strategische doelstelling is, dan zijn deze organisaties toch meegenomen in het onderzoek.

Uiteindelijk zijn de case study organisaties alle drie opererend in of gelieerd aan de financiële wereld, maar elke organisatie biedt andere diensten en producten aan en er is geen sprake van directe concurrentie.

Alle geïnterviewden voldeden aan de gewenste senioriteit en werkten minimaal één jaar bij de organisatie. De meeste geïnterviewden hebben drie of meer jaar aan data analytics ervaring en de rest van de geïnterviewden hebben een dusdanige positie in de organisatie dat er wel gesteld kan worden dat er voldoende zicht en kennis op data analytics gebied is.

De resultaten zijn geanonimiseerd en als er verwezen wordt naar een organisatie dan zal OrgA, OrgB of OrgC worden gebruikt. Verwijzingen naar geïnterviewden zal na de verwijzing van de organisatie worden geplaatst, bijv. de tweede geïnterviewde van de eerste organisatie zal als referentie OrgA_R2. Bij de eerste verwijzing naar communicatie met de geïnterviewde zal ook verwezen worden naar de datum van het interview. Met elke geïnterviewde heeft één interview plaatsgevonden, dus de daarop volgende verwijzingen hebben dezelfde datum en zal voor de leesbaarheid niet herhaald worden.

Het nut van een semigestructureerd onderzoek is gebleken, aangezien op sommige onderwerpen verder werd ingezoomd of belangrijk werden gevonden door de geïnterviewden. Als de nieuwe onderwerpen van belang werden geacht voor het onderzoek, dan werd het onderwerp ook aangehaald in andere interviews. Dit gebeurde wel meestal op het eind van het interview om het onderzoek niet te veel een richting op te sturen.

De vragen zijn uiteindelijk niet gedeeld met de geïnterviewden. De geïnterviewden zijn wel uitgebreid geïnformeerd over het onderwerp, maar het vastleggen van de eerste gedachte leidt waarschijnlijk naar minder ‘politiek correcte’ antwoorden, waardoor de resultaten beter met de werkelijkheid overeenstemmen.

De interviews zijn middels audio-opnamen opgenomen en later getranscribeerd. Oorspronkelijk was de bedoeling om voice recognition software te gebruiken, maar de gebruikte software maakte erg veel fouten, waardoor het transcriberen handmatig en door een professionele partij is uitgevoerd. De resultaten van het transcriberen zijn uiteraard wel gecontroleerd en gecorrigeerd indien nodig.

Nadat de teksten waren getranscribeerd zijn de teksten in Excel geplakt, zodat verdere analyse plaats kon vinden (Ose, 2016). Aan elk antwoord is een code toegevoegd en aan de code is het onderwerp gekoppeld. Wel bleek dat in sommige antwoorden er meerdere onderwerpen werden aangehaald. Om ervoor te zorgen dat elk onderwerp de juiste code meekreeg is de tekst van het antwoord opgedeeld in meerdere delen en aan elk deel werd een relevante codering toegevoegd (Bijlage 1). Hierdoor ontstond een fijnmaziger overzicht van coderingen.

Daarna werd per code de belangrijkste uitkomsten opgeschreven in een Word-document (Bijlage 2). Er werd eerst gezocht naar overeenkomsten in de beantwoording van de vragen, daarna naar opvallende zaken en uiteindelijk naar verschillen. Het Word-document is de directe bron voor de beantwoording van de onderzoeksvragen.

4.2. De resultaten

In deze sectie zal worden ingegaan op de resultaten. Voor de leesbaarheid is gekozen om de resultaten onder te verdelen op basis van de subvragen uit sectie 1.3 .

4.2.1. Eerste subvraag

De eerste subvraag: Welke rol speelt data in het vaststellen van de data analytics strategy binnen een organisatie? Deze vraag is erop gericht om meer inzicht te krijgen in het gebruik en soort data in de organisatie.

Data analytics strategy

Alle geïnterviewden gaven aan dat de organisaties waar zij werken data erg belangrijk vinden. Bij OrgA en OrgC waren de geïnterviewden niet op de hoogte van het bestaan van een specifieke data analytics strategy, maar wel hebben de organisaties data en het gebruik van data in de bedrijfsstrategie opgenomen. “Die staat wel gewoon heel prominent in de strategie. Het is één van de belangrijkste punten waar aan gewerkt wordt.” (OrgC_R1, persoonlijke communicatie, 2 december 2022). Voorts is er bij elke organisatie een afdeling verantwoordelijk voor data analytics. “Data is belegd in IT, daar is nou eenmaal een keuze gemaakt.” (OrgA_R1, persoonlijke communicatie, 20 oktober 2022).

In de bedrijfsstrategie is het niet specifiek benoemd, maar uit de interviews kwam naar voren dat data analytics binnen elke organisatie twee doelen heeft:

1. Optimalisatie van het ‘huidige’ proces en risicobeheersing, zodat de winst geoptimaliseerd wordt.
2. Data analytics om toekomstige verbeteringen of nieuwe producten en diensten te ontwikkelen.

Voor beide doelen wordt gebruikt gemaakt van interne en externe data.

Interne en externe data

Elke organisatie verzamelt veel interne data. Enkele voorbeelden: klikgedrag op de site, NAW-gegevens, betaalgegevens en gezinssamenstelling. Volgens OrgB_R1: "...We horen bij de meeste data scientists dat ze alle data willen hebben. Want als je alle data hebt dan kun je daar patronen uit afleiden en uit die patronen kun je misschien leuke dingen doen. Dat is gedeeltelijk waar, maar dat is niet ongelimiteerd waar." (OrgB_R1, persoonlijke communicatie, 25 oktober 2022), maar niet alles wordt gebruikt vanwege de privacywetgeving. OrgB_R1: "Je zit altijd met de AVG en privacyzaken" en OrgC_R2: "En je hebt dus ook privacyregelgeving, daar worstelen wij ook gewoon heel erg mee." (OrgC_R2, persoonlijke communicatie, 8 december 2022)

Externe data wordt verzameld met als doel om de klantbeleving te verbeteren door gegevens van klanten al voor in te vullen. OrgB_R2 stelt: "...dat je zo min mogelijk gegevens hoeft te vragen, want dat is voor de klant ook het meest fijne natuurlijk." (OrgB_R2, persoonlijke communicatie, 10 november 2022). Externe data wordt ook gebruikt om klantrisico's beter in te schatten met als doel om verliezen te beperken. Enkele voorbeelden: voertuiggegevens van de RDW (voorheen genaamd Rijksdienst voor het Wegverkeer), Bureau Krediet Registratie (BKR), de Kamer van Koophandel of het Kadaster.

Tijdens de interviews kwam naar voren dat externe data vooral bij het aangaan van een klantrelatie wordt gebruikt. Bij het klantbeheer wordt veel minder externe data gebruikt. Voor het onderzoek is ervan uit gegaan dat data verkregen van externe partijen, niet zijnde de klant, altijd beschouwd wordt als externe data. Ook als de data al langere tijd in bezit is

4.2.2. Tweede subvraag

De tweede subvraag onderzoekt in hoeverre er een onderscheid wordt gemaakt tussen in- en externe data. Zoals bij de eerste subvraag al naar voren is gekomen als resultaat, maken alle drie de organisaties gebruik van in- en externe data. Al is het gebruik van externe data in eerste instantie vooral ingegeven door het verbeteren van de klantbeleving en het beperken van risico's. Bij alle drie de organisaties loopt het gebruik van interne en externe data door elkaar heen. Het werd als volgt verwoord door OrgC_R2: "...wij maken niet heel erg bewust het onderscheid tussen intern of externe data; het is gewoon allemaal data en het een is waardevol en het ander niet." (OrgC_R2, persoonlijke communicatie, 8 december 2022). Dit behoeft wellicht iets meer uitleg.

Praktisch doel

Uit elk interview kwam naar voren dat het proces begint op het moment dat een klant een aanvraag indient voor een product, het maakt niet uit welk product. De klant voert zelf gegevens in en er wordt vanuit externe bronnen data verzameld. Als de aanvraag heeft geresulteerd in een contract, dan zal voor het beheer ook gebruik worden gemaakt van de al in bezit zijnde data. Soms verrijkt men de data uit externe bronnen of vraagt men de klant om data te updaten. Dit is niet te classificeren als data exploratie, er wordt enkel gekeken naar de huidige status en om te zien of het risicoprofiel van de klant conform verwachting is. "...zo weinig mogelijk klanten door het handmatige stuk moeten." (OrgC_R2). Is dat niet het geval, dan zullen de acceptatiecriteria aangepast worden. Deze analyse en aanpassingen zullen toekomstig effect hebben, maar wordt gezien als onderdeel van data exploitatie.

Toekomstige en strategische verbeteringen

Als er dan resources over zijn, dan wordt er gekeken of er mogelijk nog toekomstige verbeteringen mogelijk zijn. Dit kan zijn een nieuw product, dienst of een toekomstige aanpassing in het proces. Dit zal in samenspraak gaan met de verantwoordelijke operationele afdeling en er zal een hypothese worden gevormd. OrgB_R2: "...dat noemen we dan opportunity finding-sessies met de business..." Op basis van die hypothese gaat de data scientist aan de slag. Het zonder hypothese op zoek gaan in de data wordt door de geïnterviewden niet zinvol geacht. Het aantal variabelen is te groot, de data is te omvangrijk. De noodzaak voor externe data ter verrijking van de interne data is er niet direct. De verzamelde interne data is omvangrijk en men verwacht dat externe data slechts in beperkte mate toegevoegde waarde heeft op dit moment.

Beperkingen in gebruik data

Er zijn wel beperkingen ten aanzien van het gebruik van externe data. Men wordt regelmatig benaderd door marktpartijen die dataproducten aanbieden, maar men is daar terughoudend in, omdat niet altijd duidelijk is welke data wordt gebruikt en hoe. Het afnemen van een dergelijk product kan leiden tot privacy of ethische problemen en potentieel tot boetes en afbreukrisico. OrgC_R2: "...als je dan wat blackbox-achtige modellen van externe partijen koopt die niet willen vertellen wat ze doen, dan is het gewoon een moeilijk verhaal."

Een andere beperking is het verkrijgen van bepaalde data. Daar zijn een aantal mogelijke redenen voor: de interne kwaliteitscommissie geeft geen goedkeuring voor het gebruik, technische aanpassingen om de data te verzamelen en soms het op zoek gaan naar een externe dataleverancier (OrgC_R2).

De wet- en regelgeving speelt een belangrijke rol in de financiële sector en dat geldt ook bij het gebruik van data. Alle partijen geven aan dat vanwege wet- en regelgeving bepaalde, met name privacygevoelige, data niet wordt gebruikt. Een gedeelte van de data kan of mag niet worden gebruikt of verzameld. Dat geldt voor zowel interne als externe data (OrgA_R1 : "...we hadden zeg maar bijvoorbeeld 500 pagina's aan regelgeving. Ik denk dat we nu 5.000 pagina's hebben aan regelgeving.").

4.2.3. Derde subvraag

De derde subvraag betreft in hoeverre er een optimum is in het gebruik van interne en externe data. Als er gestreefd wordt naar een optimum tussen het gebruik van in- en externe data dan lijkt het of er sprake is van een aansluiting met de data analytics strategy.

Verrijken interne data

De eerder genoemde tweede doelstelling, strategische verbeteringen identificeren, vindt vooral plaats op basis van de eerder verzamelde data. De verzamelde data in het kader van de eerste doelstelling kan zowel interne als externe data zijn. Wel gaf men aan dat het mogelijk is om de data verder te verrijken met externe data.

Meerdere geïnterviewden gaven aan dat de toegevoegde waarde van het gebruik van een externe variabele afneemt bij elke toevoeging. OrgB_R1 stelt: "Het is niet zo dat hoe meer data je verzamelt

en hoe hoger de variëteit van data is, dat je analyses steeds beter worden.” en OrgC_R2 noemt het “...het is een wet van afnemende meeropbrengst.” Men wil vooral de interne data als basis gebruiken eventueel verrijkt met wat externe data.

Optimum

Bij de drie organisaties wordt er gestreefd naar een optimum in de verhouding interne en externe data. Het optimum is vooral van praktische aard, aangezien het doel in eerste instantie is om de klantbeleving te verbeteren en ter voorkoming van risico's. Elk van de geïnterviewden gaf aan dat de hoeveelheid externe data afhangt van de situatie. Met name bij het aangaan van een klantrelatie wordt er relatief gezien veel externe data gebruikt. Er werd zelfs de verhouding van 90% externe en 10% interne data genoemd (OrgC_R2). Deze verhouding wijzigt na verloop van tijd. Zeker zodra de fase van contractbeheer is aangebroken, een paar geïnterviewden hadden het over een verhouding van 95/5, waarmee werd bedoeld 95% interne en 5% externe data (OrgB_R2, OrgC_R1), maar ok ene maximum van 20 % is genoemd (OrgC_R1). “...ik weet ook niet of het heel gezond is als je dan verder gaat dan die 80-20.” (OrgB_R2). Er kan gesteld worden dat de geïnterviewden de verhouding van de context laten afhangen en dat er vooral gebruik moet worden gemaakt van interne data. De context wordt gevormd door de wet- en regelgeving, de hoeveelheid interne data, betrouwbaarheid externe bronnen en de klantgerichtheid van de organisatie. Dit is overigens geen uitputtende lijst, dit zijn de elementen die besproken zijn tijdens de interviews.

Ambidexterity

Voor data exploratie wordt niet standaard tijd vrijgemaakt. Het is eerder afhankelijk van de behoefte of de context dat er data exploratie wordt uitgevoerd. OrgB_R2; “...er is altijd wel een trigger vanuit de business...”. De organisaties hebben één afdeling die zich bezighoudt met zowel data exploitatie als data exploratie. Er is geen sprake van structural ambidexterity. Om de efficiency te verbeteren wordt er een hypothese opgesteld zodat de dataset kleiner kan worden en er gericht gezocht kan worden. In dit opzicht zijn er overeenkomsten met contextuele organisational ambidexterity. Er wordt gewerkt aan de behoeften van de organisatie op dit moment en er wordt ook gekeken naar de toekomstige behoefte van de organisatie, zonder dat er een extra afdeling voor wordt opgericht (Gibson & Birkinshaw, 2004). Dit kan ook wel contextuele data ambidexterity genoemd worden. Bij alle drie de organisaties is er sprake van een dergelijke contextuele data ambidexterity en betekent concreet dat er één afdeling zowel data exploitatie als data exploratie uitvoert, dat de verhouding tussen interne en externe data wisselt en afhankelijk is van de mogelijkheden, behoeften en resources.

Bij deze deelvraag kan geconcludeerd worden dat er een optimum wordt nagestreefd, maar het optimum is afhankelijk van de context: contextuele data ambidexterity.

5. Discussie, conclusies en aanbevelingen

In dit hoofdstuk zal gekeken worden in hoeverre de beantwoording van de deelvragen uit hoofdstuk 4 leiden tot de beantwoording van de onderzoeksvraag: hoe zorgt een organisatie voor een verhouding tussen data exploitatie en data exploratie (data ambidexterity) in de data analytics strategy?

5.1. Discussie – reflectie

Strategische keuze

Zoals in 2.3.1 beschreven vergt het gebruik maken van big data analytics wel een strategische keuze van een organisatie. Bij twee organisaties was weliswaar geen data analytics strategy bekend, maar voor het gebruik van data is wel een strategische keuze gemaakt en die keuze is ook doorgevoerd in de organisatie. In het kader van dit onderzoek is ervan uitgegaan dat het hebben van een strategische doelstelling om gebruik te maken van data en het hebben van een verantwoordelijke afdeling impliceert dat er een data analytics strategy is, al is het mogelijk niet beschreven of breed gedeeld binnen de organisatie. Dit is ook in lijn met de uitkomsten van de interviews, aangezien de organisaties wel degelijk data analyse op de verkregen data uitvoeren. Dit hoeft bij andere organisaties niet zo te zijn. De verantwoordelijke afdeling hoeft niet per se data te analyseren, men kan ook enkel als doel hebben om data te verzamelen om wellicht later door te verkopen.

Onderscheid interne en externe data

Er bleek dat er in de data analytics strategy of de strategische bedrijfsdoelstellingen geen onderscheid wordt gemaakt in externe en interne data. Dat dit onderscheid minder duidelijk is, komt wellicht doordat in de financiële sector veel interne data verzameld wordt om te voldoen aan wet- en regelgeving en omdat de klanten veel interne data ter beschikking stellen, waarop analyses uitgevoerd worden. Mogelijk wordt er in sectoren met minder interne data een duidelijker onderscheid gehanteerd. Tijdens de interviews is wel overwogen om hier nog wat dieper op in te gaan, maar het gevaar zou zijn dat er antwoorden zouden worden gegeven die niet aansluiten bij de praktijk, aangezien de geïnterviewden mogelijk geen ervaring hebben in andere sectoren en er te veel aannames worden gedaan.

Contextuele data ambidexterity

Tijdens de interviews leek het gebruik van contextuele data ambidexterity een bewuste keus te zijn. Het is goed mogelijk dat door het ontbreken van een formele en breed gecommuniceerde data analytics strategy bij twee organisaties (OrgA en OrgC) gekozen is voor deze vorm, nog steeds een bewuste keuze, maar wel een keuze die mogelijk niet expliciet wordt gedragen door het senior management. Contextuele data ambidexterity is door de flexibiliteit makkelijk aan te passen aan de omstandigheden en de behoeften van de organisatie, maar daardoor is er ook een kans dat data exploratie niet wordt uitgevoerd. Tijdens de interviews had hier nog een controle vraag over kunnen worden gesteld.

Nadelen data exploratie

De nadelen van data exploratie (March, 1991) kwamen ook tijdens dit onderzoek naar voren. Men gaf aan dat het niet zeker is of en wanneer “de pot met goud” (OrgB_R2) gevonden wordt en het is lastig om de organisatie mee te krijgen met wijzigingen. Dit geldt zeker als de gevonden resultaten leiden tot het reduceren van het aantal medewerkers (OrgA_R1). Een ander nadeel van data exploratie is ook gevonden: de resultaten wijken gevoelsmatig te veel af van de werkelijkheid, waardoor de angst ontstaat om iets met de resultaten te doen. Gedurende de interviews is niet gevraagd naar mogelijke oplossingen om de angst voor nieuwe inzichten te beperken. Achteraf bezien had dit van toegevoegde waarde kunnen zijn, omdat dit relevant is voor de algemene acceptatie van de resultaten en mogelijke acties die voortvloeien uit data exploratie.

Gebruik van hypothesen bij data exploratie

Er wordt vooral gebruik gemaakt van de beschikbare data om daarop data exploitatie uit te voeren en de vragen die nu spelen beantwoorden. Dit is in lijn met ander onderzoek. Data exploratie vindt wel plaats, maar zal altijd gebeuren nadat er hypothesen zijn geformuleerd. Er wordt dan gezocht naar toekomstige verbeteringen in een bepaalde richting. Het grasduinen door de data zonder richting vindt niet of zelden plaats. De reden hiervoor is dat bij alle drie de organisaties de hoeveelheid data te groot is om binnen een redelijke tijd met bruikbare resultaten te komen. Men heeft er voor gekozen om hypothesen te formuleren voordat men de data gaat analyseren. Men hoopt hiermee gerichter te kunnen zoeken en tijd te besparen. Hier komt het spanningsveld tussen data exploitatie en data exploratie duidelijk naar voren. Het hebben van een hypothese bij de data exploratie kan mogelijk een beperkend effect hebben op het vinden van een concurrentievoordeel. In het onderzoek is het gebruik van hypothesen geschaard onder data exploratie, maar mogelijk kan het gebruik van hypothesen ook worden gebruikt als tussenvorm tussen data exploitatie en data exploratie in. Wellicht is dit iets voor verder onderzoek.

Contextbepalende factoren

Op basis van de analyse op de resultaten zijn er een aantal zaken gedestilleerd die de context bepalen: wet- en regelgeving, hoeveelheid interne data, betrouwbaarheid externe bronnen en de klantgerichtheid van de organisatie. Deze resultaten zijn niet ter bevestiging voorgelegd aan de geïnterviewden en het is mogelijk dat er nog meer elementen de context bepalen. Deze vraag had voor de volledigheid opgenomen moeten worden in de vragenlijst.

Het onderwerp 'strikte wet- en regelgeving' kwam in elk interview ter sprake. Het valt buiten het bestek van dit onderzoek om te onderzoeken of de regeldruk in de financiële wereld daadwerkelijk groter is dan in andere sectoren.

Senior management

Er is de vraag gesteld hoe belangrijk de organisatie data vindt, maar die gaat niet specifiek over het senior management. Het senior management dient ook de toegevoegde waarde van data exploratie te zien en te steunen (OrgA_R2, persoonlijke communicatie, 27 oktober 2022). Is dat niet het geval dan heeft het hebben van een geformaliseerde strategie minder toegevoegde waarde. De vraag kan namelijk ook omgedraaid worden. Stel dat er een verhouding is tussen data exploitatie en data exploratie in de data analytics strategy en het senior management heeft vooral behoefte aan antwoorden op vragen die nu spelen. Dan is de kans alsnog groot, dat data exploratie geen prioriteit krijgt. Het hebben van contextuele data ambidexterity is in dit geval een goede oplossing.

5.2. Conclusies

De onderzoeksvraag (Hoe zorgt een organisatie voor een verhouding tussen data exploitatie en exploratie (data ambidexterity) in de data analytics strategy?) is niet eenvoudig te beantwoorden. Het onderzoek richt zich op enkele financiële instellingen met een aantal gemeenschappelijke kenmerken: strikte wet- en regelgeving, veel interne data en data analyse is een onderdeel van de bedrijfsstrategie. Dit betekent niet automatisch dat de organisaties een data analytics strategy hebben. Bij twee organisaties is die data analytics strategy namelijk niet aanwezig en bij de derde organisatie is er wel een data analytics strategy aanwezig, maar daarin wordt niet voorgeschreven hoe de verhouding tussen data exploitatie en exploratie moet zijn.

Ondanks het ontbreken van een data analytics strategy hebben de organisaties gekozen voor contextuele data ambidexterity. Er is één afdeling verantwoordelijk voor de data en analysewerkzaamheden en afhankelijk van de behoefte door de stakeholders o.a. senior management, wordt er gekozen voor data exploitatie en data exploratie.

Het vastleggen van een verhouding tussen data exploitatie en data exploratie in een data analytics strategy zal zeker bijdragen aan meer focus op data exploratie, omdat er dan expliciet is bepaald dat er aan data exploratie dient te worden gedaan, maar in de praktijk sluit contextuele data ambidexterity beter aan bij de behoeften en wensen van de organisaties. Het biedt de mogelijkheid om flexibel en adequaat op de wensen van de stakeholders in te gaan. Het vastleggen van een verhouding zal die flexibiliteit beperken.

In ieder geval kan geconcludeerd worden dat, ondanks dat er een verhouding ontbreekt in de data analytics strategy en er soms zelfs geen data analytics strategy is, er toch een balans wordt gezocht die op dat moment past binnen de mogelijkheden en behoeften. Het hebben van een data analytics strategy heeft wel degelijk toegevoegde waarde voor het daadwerkelijk vrijmaken van tijd en resources voor data exploratie, maar uiteindelijk zijn er ook andere factoren, zoals de beschikbare data, de wet- en regelgeving en prioriteit van data exploratie in de organisatie, die een rol spelen bij het inzetten van resources op data exploratie en het vergroten van de kans op concurrentievoordeel. Concurrentievoordeel behalen door data exploratie is meer dan een strategie.

Voortvloeiend uit het onderzoek worden in de volgende secties enkele aanbevelingen gedaan

5.3. Aanbevelingen voor de praktijk

Om te borgen dat er aandacht wordt besteed aan data exploratie is het wel noodzakelijk om in de strategie daar ruimte voor te creëren. Dat kan in de strategisch doelstelling van de organisatie zijn, maar ook verder worden uitgewerkt in de data analytics strategy. Er moet duidelijk een onderscheid in data exploitatie en data exploratie worden gemaakt, zodat er ook een bewustzijn wordt gecreëerd in de organisatie en bij het senior management. Zeker in het geval een organisatie gebruikt maakt van contextuele data ambidexterity, dan kan het bijdragen aan een focus op data exploratie.

Het operationaliseren van de analyse uitkomsten is lastig in de praktijk, zeker als het gevoel niet overeenstemt met de uitkomsten. In de praktijk leidt dit vaak tot de situatie dat er niets met de uitkomsten wordt gedaan, wat mogelijk een gemiste kans op concurrentievoordeel is. Wellicht in dat soort gevallen, kan een interne of externe second opinion ervoor zorgen dat er toch met de resultaten rekening wordt gehouden.

Data exploratie krijgt vanwege de tijdsinvestering tezamen met een onzekere uitkomst over het algemeen minder prioriteit dan data exploitatie, wellicht dat artificial intelligence (AI) gedeeltelijk de bezwaren kan wegnemen. Een machine die zelfstandig op zoek gaat naar het “goud” in de data maakt wellicht mogelijk dat er een betere borging tussen data exploitatie en exploratie plaatsvindt.

5.4. Aanbevelingen voor verder onderzoek

Zoals eerder opgemerkt worden de wet- en regelgeving in de financiële sector als erg strikt ervaren door de geïnterviewden, wat ook zijn weerslag heeft op het gebruik van data en data analytics. Hierbij valt op te merken dat ondanks de komst van ‘fintech’ het aantal disruptieve producten, die tot mogelijk concurrentievoordeel zouden kunnen leiden, beperkt is. Het gebruik van data en data analytics leidt vooral tot procesverbeteringen en niet zo zeer tot baanbrekend nieuwe producten. Meer inzicht in de beperkende werking van alle regels op vernieuwing kan mogelijk leiden tot een

bewustwording en toename in het gebruik van nieuwe technologieën in de financiële sector en om een breder beeld te krijgen is het wenselijk om dit onderzoek ook bij minder streng gereguleerde branches uit te voeren.

Bij de drie organisaties wordt er wel gebruik gemaakt van data exploratie, maar er wordt richting aan gegeven door een hypothese te vormen samen met collega's van de verantwoordelijke operationele afdeling. Hiermee wordt ondervangen dat de data scientist veel tijd verliest in analyses zonder de betekenis van de data te kennen op zoek gaat naar verbeteringen en wordt hiermee de hoeveelheid data beperkt. Het valt buiten het bestek van het onderzoek om te onderzoeken op welke wijze data exploratie het meest effectief is voor het krijgen van een concurrentievoordeel. Wellicht dat een hypothese wel werkt voor het vinden van een relatief kleine verbetering, maar voor het verkrijgen van concurrentievoordeel zou het best kunnen zijn, dat het hebben van geen hypothese beter werkt.

In het verlengde ligt ook de constatering dat in sommige gevallen de resultaten uit de analyse terzijde worden geschoven als het niet overeenstemt met het gevoel. Dit kan ook interessant zijn om verder te onderzoeken.

Contextuele data ambidexterity is afhankelijk van de context en het kan mogelijk interessant zijn om te onderzoeken welke factoren een significante rol spelen in de context.

Het is mogelijk dat in sectoren met minder interne data en minder strikte wet- en regelgeving het aantrekkelijker is om meer prioriteit aan data exploratie te geven. Dit zal onderzocht moeten worden.

Referenties

- Benitez, J., Llorens, J., & Braojos, J. (2018). How information technology influences opportunity exploration and exploitation firm's capabilities. *Information and Management*, 55(4), 508-523.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101.
- Duan, Y., Cao, G., & Edwards, J.S. (2018). Understanding the impact of business analytics on innovation. *European Journal of Operational Research*, 281 (2020), 673–686.
- Duncan, R.B. (1976). The Ambidextrous Organization: Designing Dual Structures for Innovation. *The Management of Organization*, 1, 167-188.
- Gibson, C.B., & Birkinshaw, J. (2004). The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity. *The Academy of Management Journal*, 47 (2), 209-226
- Harlow, H.D. (2018). Developing a knowledge management strategy for data analytics and intellectual capital. *Meditari Accountancy Research*, 26 (3), 400-419.
- Heckmann, C.S., & Maedche, A. (2018). IT ambidexterity for business processes: the importance of balance. *Business Process Management Journal*, 24 (4), 862-881.
- Idreos, S., Papaemmanouil, O., & Chaudhuri, S. (2015). *Overview of Data Exploration Techniques*. SIGMOD/PODS'15: International Conference on Management of Data, 277-281.
- Junni, P., Sarala, R.M., Taras, V., & Tarba, S.Y. (2013). Organisational ambidexterity and performance: A meta analysis. *The Academy of Management Perspectives*, 27 (4), 299-312.
- Kiron, D., & Shockley, R. (2011). Creating business value with analytics. *MIT Sloan Management Review*, 53 (1), 57–63.
- Kumar, S., & Singh, V. (2018). *Understanding Data Exploration Search A brief Study of User's Exploratory Search facets*. 2018 4th International Conference on Computing Communication and Automation (ICCCA), 1-6
- LaValle, S., Lesser, E., Schockley, R., Hopkins, M.S., & Kruschwitz, N. (2011). *MIT Sloan Management Review*, Geraadpleegd op 18 juli 2022, van <https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/>
- Liang, H., Wang, N., & Xue, Y. (2022). Juggling Information Technology (IT) Exploration and Exploitation: A Proportional Balance View of IT Ambidexterity. *Information Systems Research*, 0 (0).
- March, J. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2, 71-87.
- Maritz, J., Eybers, S., & Hattingh, M. (2020). Implementation Considerations for Big Data

Analytics (BDA): A Benefit Dependency Network Approach. *International Federation for Information Processing, LNCS 12066*, pp. 481–492.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*. Geraadpleegd op 18 juli 2022, van <https://hbr.org/2012/10/big-data-the-management-revolution>

Müller, O., Fay, M., & Von Brocke, J. (2018). The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems*, 35(2), 488-509

Niebel, T., Rasel, F. & Viete, S. (2019). BIG data – BIG gains? Understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology*, 28 (3), 296 – 316.

O'Reilly, C., & Tushman, M. (2013). Organizational ambidexterity: Past, present and future. *The Academy of Management Perspectives*, 27(4), 324-338.

Okoli, C., Schabram, K. (2010). "A Guide to Conducting a Systematic Literature Review of Information Systems Research". *Sprouts: Working Papers on Information Systems*, 10 (26).

Ose, S.O. (2016). Using Excel and Word to Structure Qualitative Data. *Journal of Applied Social Science*, 10(2), 1-16.

Popovič, A., Hackney, R., Tassabehji, R. & Castelli M. (2016). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, 20, 209–222.

Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M.L. (2009). Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance. *Organization Science*, 20(4), 685-695.

Rialti, R., Marzi, G., Caputo, A. & Mayah, K.A. (2020). Achieving strategic flexibility in the era of big data: The importance of knowledge management and ambidexterity. *Management Decision*, 58(8), 1585 – 1600.

Tsai, C., Lai, C., Chao, H., & Vasilakos, A.V. (2015). Big data analytics: a survey. *Journal of Big Data*, 2(21), 1 – 32.

Wang, Y., Kung, L. & Byrd, T.A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting & Social Change*, 126, 3–13.

Bijlage 1. Codes

Onderwerp/thema/steekwoord	Code
Geïnterviewde	0
Organisatie	1
Branche	2
Toeziçthouder	3
Historische data	4
Belang van data	5
Data analytics	6
Data analytics strategy	7
Data analytics verantwoordelijke	8
Operationalisatie van analyse uitkomsten	9
Soorten data	10
Onderscheid interne en externe data	11
Interne en externe data in de data analytics strategy	12
Gebruik van data in de strategie	13
Optimum interne en externe data	14
Het borgen van een optimum interne en externe data	15
Analyse resources	16
Gebruik van data in de branche	17
Inzet vrije uren voor data exploratie	18
Gebruik hypothesen	19
Data lakes	20
Single sources of truth	21
Data vs gevoel	22
Verrijken met externe data	23
Testen van uitkomsten	24
Invloed van concurrentie	25
Data analytics strategy niet expliciet	26
Data exploratie als neveneffect	27
Privacy	28
Klanten	29
Klantvriendelijke manier van verzamelen data	30
Senior management lost eerst korte termijn vragen op	31

Bijlage 2. Analyseresultaat

Als eerste wordt de code weergegeven en daarachter de categorie.

Code 0. Geïnterviewde

Deze code is gebruikt voor informatie over de geïnterviewde.

Code 1. Organisatie

Deze code is gebruikt voor informatie over de organisatie.

Code 2. Branche

Het onderzoek is gehouden bij drie financiële instellingen. Elk van de instellingen biedt andere diensten en opereren in een sterk gereguleerde omgeving.

Code 3. Toezichthouder

Als uitdaging werd door alle geïnterviewden de regeldruk door de wet- en regelgeving genoemd. Enerzijds vanwege steeds meer regels waar aan voldoen moet worden wat uiteindelijk ook tijd kost en anderzijds doordat niet alle data gebruikt kan worden vanwege de privacy regels.

Code 4. Historische data

Eén organisatie benoemde echt historische data, maar alle organisaties maken uiteindelijk gebruik van 'historische' data. Dat kan in- en extern zijn.

Code 5. Belang van data

In alle drie de organisaties wordt data erg belangrijk gevonden. Dit wordt ook uitgedragen door het senior management.

Code 6. Data analytics

Voortvloeiend uit de belangrijkheid van data is ook data analyse erg belangrijk in alle drie de organisaties. Er worden meerdere doelen genoemd: gebruik van data in het kader van het voldoen aan wet- en regelgeving, verbetering van het proces, kostenreductie, winstoptimalisatie en nieuw producten.

Code 7. Data analytics strategy

Bij elke organisatie is het gebruik van data en data analyse een onderdeel van de bedrijfsstrategie. Twee van de drie organisaties hebben geen data analytics strategy of hebben die strategie in ieder geval niet kenbaar gemaakt. Dat betekent niet dat de organisatie niet aan data analyse doet. Elke organisatie heeft een data science afdeling opgericht.

Code 8. Data analytics verantwoordelijke

Bij één organisatie heeft IT de analysecapaciteit. Bij de andere twee organisaties zijn er aparte analyse-afdelingen opgericht.

Code 9. Operationalisatie van analyse uitkomsten

Het in de praktijk implementeren van de uitkomsten van de analyses is bij alle drie de organisaties erg lastig. De genoemde redenen zijn:

- De resultaten worden niet vertrouwd.
- De bereidheid om te veranderen heeft ook te maken of er sprake is van een mogelijk onnodig worden van personeel.
- Het ontbreekt aan creativiteit bij de medewerkers om de analyseresultaten om te zetten tot een concreet en productief idee.

Het initiatief zal vooral vanuit het senior management moeten komen. Een goede basis biedt een strategie.

Een manier om te operationaliseren is het organiseren van ideation workshops. Medewerkers met een verschillende achtergrond samenbrengen om de resultaten van de data analyse te interpreteren en te bedenken wat er mee gedaan kan worden. Dat wordt alle drie de organisaties op de een of andere manier gedaan.

Code 10. Soorten data

De data die door de organisaties verzameld wordt is divers. De doelen van de dataverzameling kunnen als volgt geïnclassificeerd worden:

1. Voldoen aan wet- en regelgeving
2. Meer klantinformatie verkrijgen om de klant beter te helpen
3. Optimaliseren bedrijfsresultaat (betere inschatting maken van risico's)

De verkregen data dient over het algemeen een direct doel. Het verkrijgen van data om data exploratie uit te voeren is geen doel op zich. Het is een bijeffect.

Code 11. Onderscheid interne en externe data

Bij alle drie de organisaties wordt externe data gebruikt om de klant minder te laten invullen, om de datakwaliteit te verbeteren en om betere beslissingen nemen en om risico's te beperken.

Code 12. Interne en externe data in de data analytics strategy

Het verkrijgen van externe data is niet vastgelegd in de strategie. Externe data wordt vooral verkregen met het oog op het uitvoeren van de huidige processen. Er wordt niet specifiek voorgeschreven dat externe data verkregen moet worden.

Code 13. Gebruik van data in de strategie

De doelen beschreven in de data analytics strategy zijn vooral geschreven met in het achterhoofd de interne data. Ook een verdere uitbreiding van het verzamelen van interne data wordt ook genoemd.

Code 14. Optimum interne en externe data

Een optimum is moeilijk aan te geven. De geïnterviewden geven vooral aan dat een optimum niet vastligt. Veel interne data en externe data heeft minder toegevoegde waarde.

De behoefte van externe data is soms ook afhankelijk van het moment, bijv. bij acceptatie meer dan bij beheer. Dit heeft ook te maken met het feit dat de focus ontbreekt op externe data.

Code 15. Het borgen van een optimum interne en externe data

Dat is een moeilijke vraag. Een geïnterviewde gaf aan dat het in stappen zou gaan. Eerst uitputtende de interne data (beschikbare) gebruiken en dan kijken waar er nog verbeteringen mogelijk zijn.

Het optimum is afhankelijk van het soort branche en organisatie. Ook zijn er een aantal kernvoorwaarden van toepassing zoals: 1. het hebben van een solide technologische setup. Sterke IT infrastructuur, helder en robuust en 2. het hebben van de juiste mensen.

Een andere geïnterviewde liet het optimum afhangen van de toegevoegde waarde. Niet altijd betekent het hebben van meer data betere resultaten.

De vijfde geïnterviewde gaf aan dat het afhankelijk is van de omstandigheden. Wel heeft de organisatie erg veel interen data, dus de toegevoegde waarde van externe data zal minimaal zijn.

De zesde geïnterviewde gaf aan dat een optimum bereiken vanzelf gaat. Soms dan kost het wat meer tijd om extra data te verkrijgen, maar het bevordert wel de goede ideeën.

Code 16. Analyse resources

De resources zouden niet bij IT moeten zijn, maar bij de business en functionele eigenaren.

De pool van resources zouden flexibel ingezet moeten kunnen worden op basis van de behoeften van de markt. Te denken valt aan een resource management systeem.

En senior management dient achter de inzet van resources te staan.

Code 17. Gebruik van data in de branche

In de leasemarkt is het klantperspectief belangrijk. En als tweede perspectief is hoe de waarde van de voertuigen over 3 tot 5 zal zijn.

In de verzekeringsbranche wordt het gebruikt om: te voldoen aan wet- en regelgeving, privacy en om goedkoper te zijn dan de rest. Men wil zo veel mogelijk data verzamelen, uiteraard op een ethische wijze, om daarna positieve impact te maken op het welzijn van de klanten.

Met het gebruik van data kan de klant uiteindelijk ook beter worden omdat er beter een inschatting gemaakt kan worden van de situatie van de klant. Waarna er bijv. minder rente in rekening wordt gebracht.

Code 18. Inzet vrije uren voor data exploratie

Medewerkers gaan niet zelfstandig door de data grasduinen. Ook met de agile werkwijze zal men eerder werken aan de volgende sprint dan de tijd gebruiken om te gaan grasduinen. Vooral unaniem.

Code 19. Gebruik hypothesen

De algemene gedachte van de geïnterviewden was dat er zonder hypothesen of ruwe ideeën het moeilijk is om op zoek te gaan naar interessante zaken. Meestal gebeurt het na een bepaalde aanname of iets wat niet goed loopt.

Opportunity finding session: samen met business op zoek gaan naar verbeteringen.

Code 20. Data lakes

Wordt nog niet gebruikt, al zijn er wel plannen (1 geïnterviewde)

Code 21. Single sources of truth

Dit is vooral bij de verzekeraar van toepassing. Is elementair.

Code 22. Data vs gevoel

In alle 3 de organisaties komt het voor dat de uitkomsten uit de analyse gevoelsmatig niet correct, waarna er ook niks met de uitkomsten wordt gedaan.

Code 23. Verrijken met externe data

Aandachtspunten: bij grote organisaties kunnen er op meerdere plekken dezelfde data binnenkomen.

Open insurance is een trend: data reciprociteit (jij krijgt van mij data maar ik wil er ook data voor terug). Dit kan uiteindelijk leiden tot nieuwe producten en diensten (denk aan open insurance).

De mogelijkheden van externe data kunnen beperkt zijn, zeker als er al veel interne data beschikbaar is.

Het is best wel aantrekkelijk om af en toe een pilot te doen met externe data.

De toegevoegde waarde wordt steeds minder van externe data. Hoe meer je toevoegt hoe beperkter de impact.

Code 24. Testen van uitkomsten

Testen is lastig, maar men doet wel AB-testen, fail fast en fail cheap.

Code 25. Invloed van concurrentie

Een concurrentievoordeel is afhankelijk van de markt, bijv. banken zijn door hun lange bestaan betrouwbaar en innovatie is niet noodzakelijk. Ander onderzoek?

Doordat er veel data is en ook wordt gebruikt binnen de processen, is er wel degelijk sprake van een concurrentievoordeel.

Code 26. Data analytics strategy niet expliciet

De data analytics strategy is in alle organisaties niet expliciet gemaakt. Het gebruik van data, data analyses en de resources vinden vooral plaats op basis van de algemene bedrijfsstrategie.

Code 27. Data exploratie als neveneffect

Soms zien andere afdelingen op basis van data exploratie interessante mogelijkheden, bijv. een Risk Management afdeling die ook potentiële nieuwe klantprofielen ziet.

Code 28. Privacy

Privacy is bij alle 3 organisaties erg belangrijk. Het gebruik van persoonlijke data wordt voorkomen. Wel beperkt dit het gebruik en de analysemogelijkheden.

Ook wordt hierdoor het gebruik van externe data bemoeilijkt, bijv. door assessments op de toeleverancier of zorgen dat er een veilige verbinding is etc.

Code 29. Klanten

Verschillende soorten klanten. Het gebruik van data kan ook de customer journey verbeteren, bijv. bij de aanvraag kan het aanvraagformulier voor de helft gevuld doordat er een koppeling is met externe data.

Code 30. Klantvriendelijke manier van verzamelen data

Het effortless een funnel inrichten is lastig, maar helpt zeker om klantvriendelijker te werken.

Code 31. Senior management lost eerst korte termijn vragen op

Focus op korte termijn doelstellingen is aanwezig, maar is een menselijke reactie.