

MASTER'S THESIS

Data Analytics Capabilities in het Nederlandse MKB

van Riel, D.

Award date:
2022

[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: 02. Jul. 2022

Open Universiteit
www.ou.nl



Data Analytics Capabilities in het Nederlandse MKB

Data Analytics Capabilities in Dutch SMEs

Opleiding: Open Universiteit, faculteit Betawetenschappen
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: Dylan van Riel

Identiteitsnummer:

Datum: 15-03-2022

Afstudeerbegeleider Jeroen Baijens

Meelezer Prof.dr.ir. Remco Helms

Derde beoordelaar <indien aanwezig>

Versie nummer: 1.0

Status: Final

Abstract

Het doel van deze studie is om te valideren welke dimensies bijdragen aan de opbouw van een DA Capability voor het MKB in Nederland. De focus lag hierbij op Tangible Resources, onderdeel van het conceptueel model afgeleid uit eerder onderzoek door Moonen en collega's. Voor deze studie is gebruik gemaakt van een zelfontwikkelde vragenlijst (Cronbachs $\alpha=.79$). Data van 125 respondenten is geanalyseerd door middel van Exploratieve Factor Analyse (EFA). Resultaat is dat het conceptueel model grotendeels toepasbaar is op het MKB, met als uitzondering dat minder onderliggende factoren benodigd zijn en dat binnen het MKB minder gebruik gemaakt wordt van externe data. Voor het MKB zijn twee factoren gedefinieerd die samen de dimensie Tangible Resources vormen: Data Toepassing en Data Infrastructuur. De resultaten kunnen als validatie dienen van IT dienstverlening, of als gespreksinitiator van een salesorganisatie. Met deze studie is een bijdrage geleverd aan de kennis van de opbouw van DA Capabilities voor het MKB op basis van Tangible Resources. Een aanbeveling voor vervolgonderzoek is bekijken of deze Capabilities ook een invloed hebben op bedrijfsperformance.

Sleutelbegrippen

Data, Analytics, Capability, MKB, Tangible, Resources

Samenvatting

Inzichten creëren uit data blijkt niet altijd gemakkelijk, zeker voor het MKB. Grotere organisaties gebruiken Data Analytics (DA) om deze inzichten te creëren, die uiteindelijk een competitief voordeel kunnen opleveren. Het MKB blijft hier echter bij achter. Binnen het MKB wordt DA traag geadopteerd en wordt lastig meerwaarde gecreëerd door middel van DA. De juiste capabilities zijn niet (voldoende) aanwezig om de volledige waarde uit DA te halen. Met dit onderzoek wordt getracht inzicht te bieden in het opbouwen van DA Capabilities, gefocust op de dimensie Tangible Resources. Deze dimensie wordt hierbij opgebouwd op basis van de subdimensies Data Collectie, Data Analytics, Data Architectuur en Technologie Infrastructuur. Om inzicht te bieden in de opbouw van DA Capabilities voor het MKB op basis van Tangible Resources is de volgende onderzoeksvraag gedefinieerd: Welke subdimensies vormen een valide meetinstrument voor het meten van Tangible Resources voor een DA Capability binnen het MKB?

Het MKB is een belangrijke sector in Nederland. Het feit dat er binnen deze sector weinig waarde uit data gehaald wordt, heeft dus in Nederland een grote impact. Veel Nederlandse organisaties kunnen hun voordeel halen uit de resultaten van deze studie. Ook de wetenschappelijke discipline omtrent DA Capabilities kan zich door middel van deze studie verder ontwikkelen.

Voor deze studie is gebruik gemaakt van een zelfontwikkelde vragenlijst. Deze vragenlijst is gecreëerd op basis van eerder wetenschappelijk onderzoek van onder andere Moonen en collega's (2019). De vragenlijst is afgenomen bij 125 respondenten, verzameld door gebruik te maken van een eigen netwerk aangevuld met klanten van werkgevers van studenten en (in)directe LinkedIn connecties. Uiteindelijk zijn 87 responses valide bevonden voor analyse. De overige 38 responses zijn uitgefilterd omdat hier alleen vragen over achtergrondvariabelen zijn beantwoord, of omdat niet voldoen aan het inclusie criteria 'valt onder MKB'.

Voor de steekproef van 87 valide responses zijn beschrijvende variabelen geanalyseerd, waarbij gekeken is naar functie van der respondent, bedrijfsgrootte, sector, regio en rechtsvorm.

Vervolgens is gebruik gemaakt van een exploratieve factoranalyse. Hiervoor is een Principal Component Analysis (PCA) uitgevoerd op de data. De Kaiser-Meyer-Olkin (KMO) heeft bevestigd dat het sample valide is voor analyse, KMO = .78. Bartlett's test of sphericity levert een Chi-square van 580.253, $p < .001$, wat voldoende correlatie tussen de items aantoont.

De PCA resulteert in twee factoren: Data Toepassing en Data Infrastructuur. De cumulatieve verklaarde variantie van deze 2 factoren is 49.7%. Dit in tegenstelling tot het gebruikte conceptuele model, waarbij vier subdimensies gedefinieerd zijn. Een mogelijke verklaring voor het beperkte aantal factoren ten opzichte van het conceptuele model, is dat het MKB meer gebruik maakt van standaardsoftwarepakketten voor DA. Hierdoor wordt minder fijnmazig nagedacht over de inzet van DA binnen het MKB, waardoor het onderscheid tussen bijvoorbeeld beschrijvende analyse, voorspellende analyse en voorschrijvende analyse niet gemaakt wordt. Resultaat van de uitgevoerde PCA is ook dat de variabele Externe Data onvoldoende bijdraagt aan de opbouw van de dimensie Tangible Resources. Een mogelijke verklaring hiervoor kan zijn dat het MKB onvoldoende expertise in huis heeft om externe data te extraheren, integreren met interne data en te gebruiken.

Op basis van de resultaten kan geconcludeerd worden dat alle subdimensies, meegenomen in dit onderzoek onderdeel zijn van Tangible Resources, op externe data na. Echter is de groepering van deze subdimensies wel anders dan in het conceptueel model. Dit laat dus zien dat er grote overeenkomsten zijn tussen de opbouw van een DA Capability op basis van Tangible Resources binnen het MKB en de opbouw voor grotere organisaties. Een DA Capability wordt daarbij opgebouwd door de subdimensies Data Toepassing en Data Infrastructuur.

De resultaten kunnen als validatie dienen van IT dienstverlening, of als gespreksinitiator van een salesorganisatie. Met deze studie is een bijdrage geleverd aan de kennis van de opbouw van DA Capabilities voor het MKB op basis van Tangible Resources. Een aanbeveling voor vervolgonderzoek is bekijken of deze Capabilities ook een invloed hebben op bedrijfsperformance.

Summary

To create insights from data is easier said than done, especially for SME organizations. Bigger organizations use Data Analytics (DA) to create these insights, which can create a competitive advantage. Still, SME organizations lack behind. Data Analytics is adopted much slower and thus, less added value is created by means of DA. The right capabilities are not present to create maximum added value. With this study, insights are created how DA Capabilities are created, focusing on the dimension Tangible Resources. This dimension is constructed via the subdimensions Data Collection, Data Analytics, Data Architecture and Technology Infrastructure. The research question is defined as follows: Which subdimensions create a valid measuring instrument for measuring Tangible Resources that form a DA Capability within SME organizations?

SME Organizations are a big part of the Dutch economy. Therefore, less added value created from data within this sector has a big impact. A lot of Dutch organizations can take advantage with the results of this study. Also, the scientific discipline of DA Capabilities can be developed further.

For this study, a self-developed questionnaire is used. This questionnaire is created by using previous research conducted by Moonen and colleagues (2019). The questionnaire is filled in by 125 respondents, by making use of the network of students, complemented by clients of employees and (in)direct LinkedIn connections. 87 Valid responses are used, the other 38 responses are filtered out because they do not comply with the inclusion criteria 'is SME organization'. Descriptive variables are analysed, which include role within organization, size of the organization, sector, region and legal form.

An Explorative Factor Analysis is used, by conducting a Principal Component Analysis (PCA) on the data. Kaiser-Meyer-Olkin (KMO) confirms a valid sample, $KMO=.78$. Bartlett's test of sphericity gives a Chi-square of 580.256, $p<.001$, which indicates enough correlation across items.

The PCA results in two factors: Data Application and Data Infrastructure. The cumulative explained variance is 49.7%. This in contrary to the used conceptual model, where four subdimensions are

defined. A possible explanation can be that SME makes more use of standardized software for DA. Therefore, less detailed is thought about the application of DA within the organization. For example, the distinction between descriptive analysis, predictive analysis and prescriptive analysis is often not made. A second result is that the variable External Data is contributing in a negligible extent to creating the dimension Tangible Resources. A possible explanation can be that SME organizations have less expertise on-board to extract, integrate and use external data.

Based on the results of this study, can be concluded that all subdimensions analysed are part of Tangible Resources, except for external data. The distribution of this data amongst factors is different from the distribution in the conceptual model though. Still, this shows that there are similarities between the creation of a DA Capability based on Tangible Resources within SME organizations and the creation for bigger organizations. A DA Capability is created by the subdimensions Data Application and Data Infrastructure.

The results can be a validation mechanism for IT services, or as an initiator for sales conversations. Next to that, a contribution to the knowledge regarding DA Capabilities for SME's based on Tangible Resources is made. A recommendation for further research is to study for example whether these DA Capabilities have an effect on firm performance.

Inhoudsopgave

Introductie	1
Aanleiding.....	1
Probleemstelling	2
Onderzoeksvraag.....	2
Relevantie	3
Aanpak in hoofdlijnen.....	3
Literatuurstudie	4
Aanpak en uitvoering	4
Resultaten.....	6
Doel vervolgonderzoek.....	12
Methoden	13
Conceptueel ontwerp	13
Technisch ontwerp.....	14
Gegevensanalyse.....	16
Reflectie t.a.v. validiteit, betrouwbaarheid en ethische aspecten.....	17
Resultaten.....	18
Data inspectie.....	18
Descriptives.....	20
Factoranalyse	23
Discussie, conclusies en aanbevelingen.....	27

Discussie – reflectie	27
Conclusies	31
Aanbevelingen voor de praktijk.....	32
Aanbevelingen voor verder onderzoek	32
Literatuur	33
Bijlagen	35
Bijlage A: Pattern Matrices.....	35
Bijlage B: Pattern Matrices Getransformeerd/Ongetransformeerd	36

Introductie

Aanleiding

‘Data is het nieuwe goud’ is een uitspraak die we steeds vaker horen. Maar om uit deze data ook inzichten te creëren en op die manier dat goud te benutten, blijkt vaak niet zo gemakkelijk (Baransel & Baransel, 2012). Het verkrijgen van inzichten en daarmee het analyseren, voorspellen en beheersen van processen binnen organisaties wordt Data Analytics (DA) genoemd (Coleman et al., 2016). Door middel van het toepassen van DA kunnen organisaties betere beslissingen nemen, die uiteindelijk een competitief voordeel kunnen opleveren (Ramadan, Shuqqo, Qtaishat, Asmar, & Salah, 2020). Volgens Gupta en George (2016) zijn DA Capabilities benodigd om DA goed in te kunnen zetten om vervolgens de bedrijfsprestaties te laten stijgen (Gupta & George, 2016). Capabilities worden opgebouwd door het combineren en uitrollen van diverse financiële, fysieke, menselijke en organisatorische resources (Bharadwaj, 2000).

Veel organisaties gebruiken DA om hun klanten beter te bedienen of processen te verbeteren. Denk hierbij bijvoorbeeld aan Amazon, wat op basis van online gedrag nieuwe producten aan hun klanten voorstelt (Huang et al., 2020). Een ander voorbeeld is Netflix, wat op basis van kijkgedrag van consumenten voorspelt welke nieuwe series en films interessant zijn voor de consument (Shih, Kaufman, & Spinola, 2007).

De meeste voorbeelden van succesvolle inzet van DA zien we binnen grote bedrijven. Maar niet alleen voor grote bedrijven is DA interessant, ook binnen het MKB liggen er kansen (Coleman et al., 2016). Deze kansen blijven echter vaak onbenut. Binnen het MKB wordt DA traag geadopteerd en wordt er lastig meerwaarde gecreëerd (Coleman et al., 2016). De juiste DA Capabilities zijn niet (voldoende) aanwezig om de volledige waarde uit DA te halen (Moonen, Baijens, Ebrahim, & Helms, 2019). Dit onderzoek is bedoeld om inzicht te bieden in hoe DA Capabilities voor het MKB opgebouwd kunnen worden en om daarnaast handvaten te bieden aan het MKB hoe meerwaarde uit DA gemakkelijker gecreëerd kan worden.

Probleemstelling

In het onderzoek van Moonen et al. (2019) is een raamwerk voor het MKB ontwikkeld op basis van diverse dimensies, gevonden in de literatuur. Binnen dit raamwerk wordt onderscheid gemaakt tussen vier dimensies, te weten Tangible Resources, Intangible Resources, Governance en Strategy. Deze dimensies worden als onderdeel van DA Capabilities voor het MKB gezien. Veel literatuur zich focust op frameworks voor grotere organisaties. Echter, het probleem is dat het MKB deze dimensies niet één op één kan gebruiken (Moonen et al., 2019). Hier is het framework van Moonen et al. (2019) voor ontwikkeld, echter is het niet duidelijk of én hoe de verschillende dimensies het succes van DA binnen het MKB beïnvloeden.

Specifiek voor de Tangible Resources bestaat een uitdaging binnen het MKB. Zo ontbreekt het aan grote budgetten voor het inhuren van technisch specialisten, is IT infrastructuur vaak verouderd en leven databronnen binnen de organisatie gefragmenteerd (Coleman et al., 2016; Moonen et al., 2019). Om waarde te creëren of verhogen door middel van DA, zullen op zijn minst Tangible Resources aanwezig moeten zijn, echter is het effect van de inzet van deze resources op de bedrijfsperformance lastig vooraf te meten (Coleman et al., 2016). Voor het MKB is deze uitdaging extra groot, aangezien kennis en geld in zeer beperkte mate aanwezig is (Coleman et al., 2016). Het is daarom niet helder hoe DA Capabilities opgebouwd worden en welke Tangible Resources hiervoor nodig zijn binnen het MKB.

Onderzoeksvraag

In deze studie wordt specifiek onderzoek gedaan naar hoe DA Capabilities opgebouwd kunnen worden op basis van Tangible Resources. De onderzoeksvraag voor dit onderzoek luidt daarom:

Welke subdimensies vormen een valide meetinstrument voor het meten van Tangible Resources voor een DA Capability binnen het MKB?

Om de hoofdvraag te kunnen beantwoorden, is allereerst literatuuronderzoek uitgevoerd. Hiermee wordt een theoretisch kader gecreëerd, wat de basis is voor het kwantitatief onderzoek. Met dit onderzoek

wordt getracht de hoofdvraag te beantwoorden en dus te bekijken of de gevonden subdimensies van Tangible Resources daadwerkelijk bijdragen aan de opbouw van DA Capabilities, voor het Nederlandse MKB.

Relevantie

Het MKB is een belangrijke sector binnen Nederland. Volgens het Jaarbericht Staat van het MKB van 2019 wordt aangegeven dat het MKB in Nederland gezamenlijk goed is voor 70.6% van de werkgelegenheid en 61.8% van de toegevoegde waarde (Nederlands Comité voor Ondernemerschap, 2019). Het feit dat er binnen deze sector weinig waarde uit data gehaald wordt, heeft dus in Nederland een grote impact. Veel Nederlandse organisaties kunnen hun voordeel halen uit de resultaten van deze studie. Door deze studie wordt gepoogd om MKB handvaten te bieden om meer waarde uit data te halen. Iets waar de gehele sector baat bij kan hebben, waar dus de gehele economie in Nederland van kan profiteren.

Ook binnen de wetenschap kan nog een stap gezet worden in het onderzoek naar de meerwaarde van DA binnen het MKB. Veel onderzoek focust op grotere organisaties, het MKB blijft hier vaak onderbelicht. Hierdoor ontbreekt een goed meetinstrument om de DA Capabilities van het MKB te meten. Moonen en collega's (2019) hebben een eerste stap gezet door een raamwerk te ontwikkelen, deze studie focust zich op het valideren van dit raamwerk voor wat betreft de Tangible Resources (Moonen et al., 2019). Op deze manier kan de discipline omtrent DA Capabilities zich verder ontwikkelen.

Aanpak in hoofdlijnen

Allereerst wordt voor dit onderzoek een literatuurstudie uitgevoerd. Deze literatuurstudie wordt uitgevoerd om een theoretisch kader te creëren, waarmee verwachtingen over Data Analytics Capabilities in het Nederlandse in het MKB gecreëerd kunnen worden. Deze literatuurstudie is richtinggevend aan het empirisch gedeelte van het onderzoek; de kwantitatieve studie.

Voor deze kwantitatieve studie wordt data verzameld van een honderdtal bedrijven uit het MKB segment. Van deze bedrijven wordt het raamwerk van Moonen et al. (2019) getoetst door middel van het

afnemen van een vragenlijst, zoals beschreven in Saunders et al. (2016) (Saunders, Lewis, & Thornhill, 2016).

Literatuurstudie

In dit hoofdstuk vindt u het theoretisch kader. Dit hoofdstuk bestaat achtereenvolgens uit de onderdelen aanpak, uitvoering, resultaten en doel van het vervolgonderzoek.

Aanpak en uitvoering

Om erachter te komen welke subdimensies binnen de dimensie Tangible Resources gedefinieerd worden en hoe de inzet van deze dimensies gemeten kan worden, is literatuuronderzoek uitgevoerd. Voor dit literatuuronderzoek is de methode voor een systematische literatuurreview gebruikt zoals beschreven door Okoli en Schabram (Okoli & Schabram, 2009). Door deze methode aan te houden wordt geborgd dat een gedegen aanpak wordt gehanteerd, die tevens transparant en reproduceerbaar is, waarmee geen literatuur wordt overgeslagen. Hieronder wordt toegelicht hoe de verschillende stappen van Okoli & Schabram zijn uitgevoerd voor dit onderzoek.

Identify the purpose. Door het doel van het onderzoek te beschrijven, kan doelgerichter te werk worden gegaan tijdens de vervolgstappen. Het doel van dit onderzoek is het opstellen van een theoretisch kader omtrent Tangible Resources en DA Capabilities en definiëren wat Tangible Resources voor het MKB inhouden.

Draft protocol and train the team. Training is voor dit onderzoek niet van toepassing, aangezien bij dit literatuuronderzoek geen vaardigheden benodigd zijn buiten de al aanwezige vaardigheden bij de onderzoeker. Het protocol voor literatuuronderzoek is de aanpak zoals in deze paragraaf beschreven.

Apply practical screen. Alleen peer-reviewed artikelen na 2010 gepubliceerd worden meegenomen. Voor deze tijd werd niet veel onderzoek naar DA gedaan en dus zijn artikelen voor deze

periode minder relevant. Daarnaast zijn alleen artikelen geschreven in het Nederlands of Engels meegenomen, in verband met de leesbaarheid voor de onderzoeker.

Search for literature. Web of Science is gebruikt om te zoeken naar literatuur. In deze database komen meerdere onderzoeksdomeinen samen, waardoor deze database de grootste set aan relevante literatuur bevat voor dit onderzoek. Twee zoekqueries zijn gebruikt voor het literatuuronderzoek. De eerste query geeft resultaten die ingaan op DA Capabilities binnen organisaties, en geeft 254 resultaten. De tweede query gaat specifiek in op de subdimensies van Tangible resources binnen de context van DA Capabilities binnen organisaties. Deze query geeft 155 resultaten. Beide queries zijn gecombineerd, waarna duplicaten zijn verwijderd. Het resultaat van deze gecombineerde, gefilterde query is 368 artikelen. De gebruikte queries zijn hieronder in Tabel 1 te zien, waarbij de afkortingen TS, TI en PY voor respectievelijk Topic, Title en Year Published staan.

Tabel 1

Queries Inclusief Zoekresultaten

#	Query	Resultaten
1	TS=("Capabilit*") AND TI=("Data" OR "Data Analytic*" OR "Business Analytic*" OR "Data Science" OR "Business Intelligence") AND TI=("Business" OR "Enterprise" OR "Compan*" OR "SME*" OR "Organization*") AND PY = 2010-2021	254
2	TS=("Capabilit*") AND TI=("Data" OR "Big Data" OR "Data Analytic*" OR "Business Analytic*" OR "Data Science" OR "Business Intelligence") AND TS=("Business" OR "Enterprise" OR "Compan*" OR "SME*" OR "Organization*") AND TI=("Data Collection" OR "Data Analytic*" OR "Data Architecture" OR "Technology Infrastructure") AND PY = 2010-2021	155
3	Result in #1 OR #2	368

Extract data. Het resultaat van de twee gecombineerde queries is 409 artikelen. Deze artikelen zijn ontdebeld, waarna 368 artikelen zijn overgebleven. Deze 368 artikelen zijn geëxporteerd en deze export bevat Titel, Auteurs, Jaartal en andere mogelijk relevante informatie.

Appraise quality. De eerste filtering van artikelen is gedaan op basis van titel. Na deze filtering is een selectie van 36 artikelen overgebleven, en zijn 332 niet bruikbaar gevonden. De oorzaken van deze niet-bruikbaarheid lopen uiteen van een ander onderzoeksgebied (GDPR binnen DA) tot een te nauw

onderwerp wat niet per se met Tangible resources te maken heeft (zoals reacties van medewerkers op DA innovaties). Door deze filtering zijn artikelen die niet relevant zijn voor het onderzoek, uitgesloten. Van de 32 overgebleven artikelen is het abstract doorgenomen, en is nogmaals een filtering gedaan op bruikbaarheid. Enkele artikelen bleken toch niet relevant voor het onderzoek. Artikelen die in deze filtering zijn uitgesloten, focussen zich na het lezen van het abstract toch op andere zaken, zoals moderende variabelen, bijvoorbeeld ‘omgeving’, of zijn toch gefocust op bijvoorbeeld alleen governance. Na deze filtering is een selectie van 16 artikelen overgebleven die relevant zijn voor het onderzoek. Deze zijn dan ook gebruikt voor dit literatuuronderzoek.

Synthesize studies. Gebaseerd op inhoud zijn de artikelen gegroepeerd, waarbij een artikel van meerdere groepen deel uit kan maken. Elke groep beschrijft een onderwerp in het theoretisch kader. De onderwerpen die hierbij behandeld worden, zijn Data Analytics Capabilities en Performance, Data Analytics Capabilities en hun karakteristieken, MKB en Tangible Resources.

Write the review. In de volgende paragraaf worden de resultaten van de review gesproken.

Resultaten

DA Capabilities

Er worden binnen de literatuur verschillende definities van DA Capabilities gebruikt. Zo omschrijven Mikalef en collega's (2018) een DA Capability als de bekwaamheid van een organisatie om strategisch en operationeel inzicht te creëren door het gebruik van data (Mikalef et al., 2018). Gupta en collega's (2016) omschrijven een DA Capability als de juiste combinatie resources die binnen een organisatie aanwezig zijn. (Gupta & George, 2016). Hierbij wordt binnen deze resources onderscheid gemaakt tussen Tangible Resources, Human Skills en Intangible Resources. In dit onderzoek wordt een DA Capability gedefinieerd als ‘De bekwaamheid van een organisatie om inzicht te creëren door het gebruik van data, bestaande uit de juiste combinatie resources aanwezig in de organisatie’.

Opbouw van DA Capabilites

Naast de drie resources gedefinieerd door Gupta en collega's (2016), zien we nog meer mogelijke interpretaties van de typen dimensies in de literatuur. Aktar en collega's (2016) spreken over Management, Technology en Talent (Aktar et al., 2016). Hier kan een parallel getrokken naar de drie dimensies van Gupta en collega's (2016), waarbij Management overeen komt met Intangible Resources, Technologie met Tangible Resources en Talent met Human Skills. Wamba en collega's (2017) praten over Infrastructure Flexibility (Tangible Resources), Management Capabilities (Intangible Resources) en Personnel Expertise Capabilities (Human Skills) (Wamba et al., 2017).

Echter zien we ook literatuur die met een andere invalshoek naar de dimensies kijkt. Zo definiëren Mikalef en collega's (2018) de subdimensies Data, Process, Technology, Organization, People en Context. Deze 6 subdimensies zijn echter uiteindelijk wel weer te scharen onder Tangible Resources, Intangible Resources en Human Skills. Een andere invalshoek van dimensies die gebruikt wordt, is het gebruik van de dimensies Integratie, Technologie/Data, Process en Management (L. Chen & Nath, 2018). Ook het gebruik van de dimensies Governance, Culture, Technology en People zien we terug (Cosic, Shanks, & Maynard, 2015).

Moonen en collega's (2019) hebben bovenstaande literatuur gebruikt voor het definiëren van het raamwerk, specifiek voor het MKB. De dimensies die hier gedefinieerd worden, zijn Tangible Resources, Intangible Resources, Governance en Strategie. Deze zullen in dit vervolgonderzoek ook gebruikt worden, om op deze manier het raamwerk van Moonen en collega's te toetsen. Alle hiervoor genoemde dimensies zijn te vinden in Tabel 2.

Tabel 2*Interpretatie Dimensies In Literatuur*

Artikel	Dimensie 1	Dimensie 2	Dimensie 3
Gupta et al.	Tangible Resources	Intangible Resources	Human Skills
Aktar et al.	Technology	Management	Talent
Wamba et al.	Infrastructure Flexibility	Management Capabilities	Personnel Expertise Capabilities
Mikalef et al.	Data / Technology	Process / Organization / Context	People
Chen et al.	Integratie - Technologie/Data	Process	Management
Cosiz et al.	Technology	Governance	People - Culture
Moonen et al.	Tangible Resources	Governance / Strategie	Intangible Resources

Noot: Een aantal artikelen gebruiken meer dan drie dimensies. Voor deze artikelen zijn meerdere dimensies samengevoegd, gescheiden door een ‘/’-teken.

Tangible Resources

Tangible Resources worden in verschillende studies op verschillende manieren gedefinieerd en gemeten. Overkoepelend gaat het om tastbare componenten die gekocht en verkocht kunnen worden. Technologie wordt vaak gezien wordt als onderdeel van deze Tangible Resources, ook worden Infrastructuur en Data genoemd als dimensies die we kunnen scharen onder de Tangible Resources (Y. Chen & Lin, 2021; Gupta & George, 2016; Mikalef et al., 2018; Wamba et al., 2017). De interpretatie van Moonen en collega's (2019) definieert vier subdimensies van Tangible Resources specifiek voor het MKB, te weten Data Collectie, Data Analytics, Data Architectuur en Technologie Infrastructuur. Dit zijn dan ook de vier subdimensies waarop gefocust wordt in dit onderzoek.

Zo ver bekend, is er nog geen valide en betrouwbaar meetinstrument om Tangible Resources te meten. Vele studies hebben gebruik gemaakt van eigen ontworpen vragenlijsten die gebaseerd zijn op de literatuur (o.a. Akter et al., 2016; Ashrafi, Ravasan, Trkman, & Afshari, 2019; Aydiner, Tatoglu, Bayraktar, Zaim, & Delen, 2019; Chatterjee, Rana, & Dwivedi, 2021; L. Chen & Nath, 2018; Gupta & George, 2016), waarbij ze betrouwbaarheid hebben getracht te borgen door middel van expertpanels (o.a. Aydiner et al., 2019; Chatterjee et al., 2021), test-hertesten (o.a. Ashrafi et al., 2019) en pilots (o.a. Chatterjee et al., 2021; L. Chen & Nath, 2018). Meerdere studies (o.a. Mikalef et al., 2018; Su et al., 2021) hebben de constructen van Gupta en George (Gupta & George, 2016) als basis genomen bij het opstellen van hun vragenlijsten.

Het overgrote gedeelte van bovengenoemde studies heeft het onderzoek gericht op grote organisaties, waarbij gekeken is of, en in welke mate, Tangible Resources onderdeel uitmaken van DA Capabilities. Deze studies hebben allen aangetoond dat dit inderdaad het geval is. Ook voor de verschillende subdimensies Data Collectie, Data Analytics, Data Architecture en Technology Infrastructure is aangetoond dat deze onderdeel uitmaken van DA Capabilities.

Bij deze studies zien we dat de orde van grootte van de verklaarde variantie binnen het construct Tangible Resources verschillend is tussen de subdimensies. Echter, wordt er aangeraden om alle constructen evenveel aandacht te besteden in zowel de praktijk als onderzoek (Akter et al., 2016).

De constatering dat Tangible Resources onderdeel uitmaken van DA Capabilities, wordt op verschillende manieren verklaard. Zo wordt beargumenteerd dat de inzet van DA Capabilities resource intensief is, waarmee bedoeld wordt dat DA Capabilities erg leunen op Tangible Resources (L. Chen & Nath, 2018). Een andere verklaring is dat Tangible Resources randvoorwaardelijk zijn voor het opzetten van DA Capabilities, zonder data en technologie om deze data te verwerken en hier analyses op los te laten is immers geen waarde uit deze data te halen (Akter et al., 2016; Ashrafi et al., 2019).

Ook binnen het MKB kan geen waarde uit data gehaald worden zonder data en technologie om deze data te verwerken. Tangible Resources zullen dus naar verwachting ook voor het MKB essentieel zijn bij het opbouwen van een DA Capability. Echter is de bijdrage van de verschillende subdimensies van Tangible Resources aan de opbouw van een DA Capability nog niet getoetst voor het Nederlandse MKB. Hiervoor is ook tijdens deze literatuurstudie ingezoomd op de verschillende subdimensies Data Collectie, Data Analytics, Data Architectuur en Technologie Infrastructuur.

Data Collectie. Onder Data Collectie verstaan we het identificeren en extraheren van databronnen en datatypen. Databronnen bestaat hier uit het vergaren van data uit interne operationele systemen, of externe bronnen. Datatypen kunnen onderscheiden worden in historisch, real-time, gestructureerd, ongestructureerd, Big Data en metadata (Moonen et al., 2019).

Verschillende studies, voornamelijk gericht op grote organisaties, hebben een positieve significante relatie aangetoond tussen Data Collectie en DA Capabilities. Zo hebben Aydiner en collega's (2019) dat er een positieve significante relatie is tussen 'Data Acquisition and Processing', wat gezien kan worden als Data Collectie, en DA Capabilities. Daarnaast tonen Gupta en George (2016) ook aan dat Data Collectie een positieve significante relatie heeft met DA Capabilities. In deze studies wordt de mate van Data Collectie gemeten, door te bekijken of data uit interne en externe bronnen wordt geëxtraheerd. Ook wordt bekeken welke typen data binnen de organisatie geïdentificeerd kunnen worden.

De verwachting is dat binnen het MKB de Data Collectie anders worden vormgegeven dan bij grotere organisaties. De assumptie is dat minder databronnen aanwezig zijn, en dat deze databronnen voornamelijk intern aanwezig zijn. Echter is de verwachting dat ook binnen het MKB Data Collectie toch randvoorwaardelijk is voor het opzetten van een DA Capability.

Data Analytics. Onder Data Analytics verstaan we het creëren van informatie uit data (Comuzzi & Patel, 2016). Data Analytics wordt ingedeeld in drie verschillende typen, te weten Descriptive (wat gebeurt er nu?), Predictive (wat gaat er in de toekomst gebeuren?) en Prescriptive (wat moeten we veranderen om de toekomst te beïnvloeden?). De subdimensie Data Analytics omvat ook het gebruik van tooling, waarbij onderscheid gemaakt wordt tussen Statistische en data mining software, en reports, dashboards, scorecards, OLAP kubussen en visualisatietechnologieën (Moonen et al., 2019).

Eerdere studies, gericht op grote organisaties, hebben modellen getoetst waarbij de subdimensie Data Analytics onderdeel is van de DA Capabilities. Zo tonen Chen en Nath (2018) aan dat Data Analytics onderdeel is van het succes van DA. Daarnaast tonen Aydiner en collega's (2019) aan dat er een positieve significantie relatie is tussen verschillende subdimensies van Data Analytics (Prescriptive Analytics,

Predictive Analytics en Descriptive Analytics) en DA Capabilities, waarbij deze subdimensies gemeten worden door de mate van inzet van deze typen Data Analytics.

De verwachting is dat de mate van Data Analytics binnen het MKB een zwakkere factorlading heeft op de opbouw van DA Capabilities, dan bij grotere bedrijven. Zo zullen veel Analytics producten zich richten op Descriptieve statistieken en zal van Predictieve en Prescriptieve statistieken weinig tot geen gebruik gemaakt worden. Echter is de inzet van Descriptieve statistieken, en daarmee Data Analytics als subdimensie, ook binnen het MKB randvoorwaardelijk voor het opzetten van een DA Capability.

Data Architectuur. Met Data Architectuur wordt het gehele proces van het opslaan, verwerken en integreren van data bedoeld (Coleman et al., 2016). Moonen en collega's (2019) verdelen Data Architectuur in Data opslag, Data verwerking, Data integratie, Data transformatie, Relatie tussen data en structuur en analyse tools.

Gupta en Collega's (2016) hebben een positieve significante relatie aangetoond tussen Data Architectuur en DA Capabilities. De subdimensie Data Architectuur wordt hier voornamelijk omschreven als Data Integratie. Hiernaast hebben Akter en collega's (2016) dat BDA Infrastructure Flexibility (opgebouwd door BDA Connectivity, BDA Compatibility en BDA Modularity) onderdeel is van DA Capabilities. Beide studies hebben onderzoek gedaan binnen grotere organisaties. Hierbij wordt Data Architectuur gemeten door te kijken naar de aanwezigheid van Data opslag, verwerking, integratie en transformatie.

Het MKB kent een kleinere verscheidenheid aan data bronnen en een verwacht groter gebruik van standaard softwarepakketten. Hierdoor speelt Data Architectuur, bestaande uit onder andere de integratie en transformatie van data, mogelijk een kleinere rol binnen het MKB. Echter wordt nog steeds wel verwacht dat Data Architectuur onderdeel is van de opbouw van DA Capabilities, immers dient de data nog steeds opgeslagen en verwerkt te worden om hier inzichten uit te creëren.

Technologie Infrastructuur. Met Technologie Infrastructuur wordt de technische datavolwassenheid van de organisatie bedoeld (Comuzzi & Patel, 2016). Binnen het raamwerk van Moonen en collega's (2019) worden Centraal Data Warehouse, Systeemintegratie en Infrastructuur beveiliging benoemd als onderdelen.

Verschillende studies gericht op grotere organisaties hebben de relatie tussen Technologie Infrastructuur en DA Capabilities onderzocht. Gupta en George (2016) hebben aangetoond dat er een positieve significante relatie bestaat tussen Technology en DA Capabilities. Hierbij wordt Technology gemeten door middel van het adopteren van parallel computing aanpakken, verschillende data visualisatie tools en open source methodieken. Vergelijkbaar met deze studie hebben Mikalef en collega's (2018) en Su en collega's (2021) met dezelfde constructen een zelfde resultaat aangetoond.

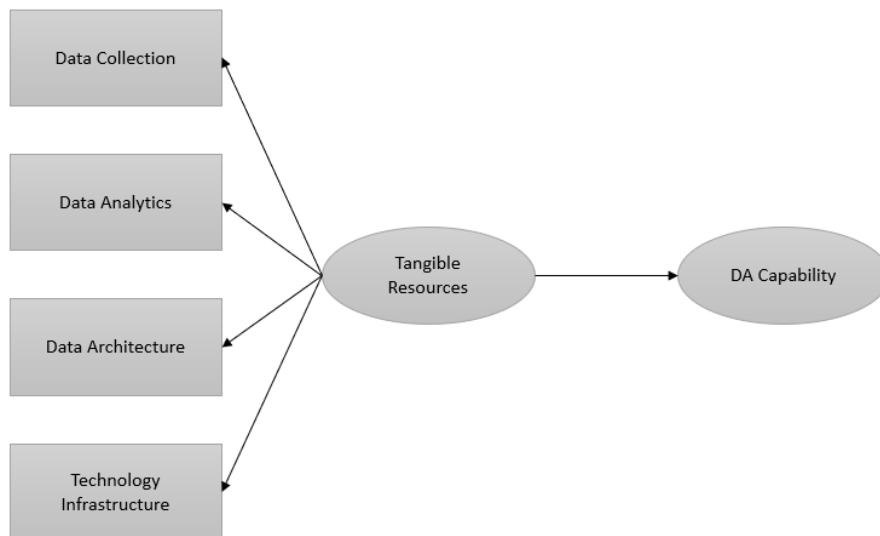
Overeenkomend met grotere organisaties zal Technologie Infrastructuur binnen het MKB naar verwachting ook onderdeel uitmaken van de opbouw van een DA Capability. Echter, zal deze subdimensie naar verwachting wel een kleinere rol spelen dan binnen grotere organisaties. Parallel computing aanpakken worden ingezet bij grote hoeveelheden data, die binnen het MKB in mindere mate aanwezig zijn. Open Source methodieken worden voornamelijk toegepast op ongestructureerde data, die binnen het MKB naar verwachting ook minder aanwezig zullen zijn. Echter zullen data visualisatie tools naar verwachting wel gebruikt (moeten) worden, zodat inzichten uit data op een begrijpelijke manier kunnen worden gepresenteerd aan eindgebruikers.

Doel vervolgonderzoek

Zoals te lezen in vorige paragraaf, hebben meerdere studies onderzoek gedaan naar DA Capabilities en de opbouw van deze DA Capabilities op basis van onder andere Tangible resources. Echter is dit onderzoek voornamelijk gedaan in grotere organisaties. Dit onderzoek focust zich daarom specifiek op het MKB, en op het toetsen van deze MKB organisaties op basis van het framework van Moonen en collega's (2019). Op basis de resultaten van de literatuurstudie kan verwacht worden dat Tangible Resources, met daarbij verschillende subdimensies, onderdeel uitmaken van DA Capabilities.

Figuur 1

Conceptueel model



Methoden

Conceptueel ontwerp

Het doel van dit onderzoek is bekijken welke subdimensies van de dimensie Tangible Resources bijdragen aan de opbouw van DA Capabilities binnen het MKB. Het raamwerk opgesteld door Moonen en collega's (2019) staat hierbij centraal. Op basis van de onderzoeksresultaten wordt bekeken of de verschillende subdimensies voor Tangible resources (Data Collection, Data Analytics, Data Architecture en Technology Infrastructure) ook daadwerkelijk laden op de dimensie Tangible Resources voor het MKB in Nederland, daarnaast wordt bekeken hoe deze subdimensies opgebouwd worden. Op deze manier wordt de bestaande theorie, het raamwerk van Moonen, getoetst. Dit onderzoek kan daarom gezien worden als deductief (Saunders, Lewis, & Thornhill, 2016).

Voor dit kwantitatieve onderzoek wordt een steekproef genomen. Deze steekproef bestaat uit MKB organisaties, waarvan vervolgens data verzameld wordt over de inzet van verschillende dimensies die door Moonen en collega's (2019) gezien worden als onderdeel van DA Capabilities. Op deze manier wordt getracht om conclusies te trekken die geldend zijn voor het MKB in Nederland. Hierbij wordt de inzet van

de DA dimensies gemeten, en de aanwezigheid van de verschillende subdimensies die deze dimensies vormen. Om een valide meetinstrument te creëren voor het meten van Tangible Resources in het Nederlandse MKB, is een objectief onderzoek nodig. Resultaten van dit onderzoek dienen met zo min mogelijk menselijke interpretatie onderzocht te worden. Dit onderzoek kan daarom geschaard worden onder het positivisme (Saunders et al., 2016).

Technisch ontwerp

Kwantitatieve data wordt verzameld van 125 MKB-organisaties, door middel van het afnemen van een vragenlijst. Deze vragenlijst bevat vragen over de aanwezigheid van de verschillende subdimensies binnen de organisaties. De subdimensies Data Collection, Data Analytics, Data Architecture en Technology Infrastructure worden gemeten voor dit onderzoek. De vragenlijst wordt gezamenlijk door 4 studenten opgesteld, en is gebaseerd op de bestaande literatuur en inzichten verkregen gedurende het literatuuronderzoek. Voordeel van deze methodiek is dat snel veel data vergaard kan worden en dat antwoorden gestandaardiseerd gegeven worden.

Zoals gemeld wordt de data vergaard door middel van het afnemen van een vragenlijst. Deze vragenlijst wordt online beschikbaar gesteld om door de respondenten ingevuld te worden. Op deze manier kan een zo groot mogelijke groep respondenten makkelijk bereikt worden, en kunnen de respondenten de vragenlijst invullen op een voor hun gunstig moment.

De functie van de vragenlijst is het vergaren van de data waarmee het kwantitatieve onderzoek uitgevoerd kan worden. Het gaat hierbij om vragen over de verschillende Data & Analytics dimensies en vragen over de bedrijfsperformance. De vragen worden beantwoord via een 5-punts Likert scale, antwoorden worden vergaard via Limesurvey. Voor een overzicht van de vragen zoals gesteld in de vragenlijst, zie Tabel 3.

Tabel 3*Overzicht items*

Nr.	Item	Vraag
B1	kanaal	N.v.t. (afgeleid uit enquêteID)
B2	directie	Wat is uw functie?
B3	bedrijfsgrootte	Wat is de grootte van uw bedrijf?
B4	sectorgroep	In welke sector is uw bedrijf actief?
B5	regio	In welke regio is uw (hoofd)kantoor gevestigd?
B6	rechtsvorm	Wat is de rechtsvorm van uw bedrijf?
A1	interneData	De organisatie maakt gebruik van interne databronnen voor data-analyse.
A2	externeData	De organisatie maakt gebruik van externe databronnen voor data-analyse.
A3	realtimeData	De organisatie maakt gebruik van real-time data voor data-analyse.
A4	gestructureerdeData	De organisatie maakt gebruik van gestructureerde data (data in tabelvorm) voor data-analyse.
A5	ongestructureerdeData	De organisatie maakt gebruik van ongestructureerde data (zoals video's, foto's of geluid) voor data-analyse.
A6	metadata	De organisatie legt metadata (beschrijvende data over andere data, zoals 'Wanneer is de data vastgelegd' en 'Wie heeft de data ingevoerd') vast en gebruikt deze voor analysedoeleinden.
A7	beschrijvendeAnalyse	De organisatie maakt gebruik van beschrijvende data-analyse om gebeurtenissen uit het verleden te begrijpen (vragen zoals 'wat is er gebeurd?').
A8	voorspellendeAnalyse	De organisatie maakt gebruik van voorspellende data-analyse zoals machine learning (vragen zoals 'wat kan er gebeuren?').
A9	voorschrijvendeAnalyse	De organisatie maakt gebruik van voorschrijvende data-analyse zoals voor procesoptimalisatie en simulatie (vragen zoals 'wat moeten we doen?').
A10	ondersteunendeToolsMiddelen	De ondersteunende IT middelen die de organisatie gebruikt om inzichten uit data te halen, voldoen aan de wensen en eisen van de organisatie.
A11	dataOpslag	De organisatie slaat data op om deze verder te verwerken en vervolgens te analyseren.
A12	dataVerwerking	De organisatie verwerkt data om deze vervolgens te analyseren.
A13	dataOpslagOplossing	De organisatie heeft een data opslag oplossing, die aansluit bij de huidige wensen en eisen van de organisatie met betrekking tot data-analyse.
A14	dataBackup	De organisatie heeft mogelijkheden tot back-up van data goed geregeld.
A15	systeemIntegratie	De organisatie heeft systeemintegratie zo ingeregeld dat deze geschikt zijn voor data-analyse.
A16	infrastructuurSecurity	De IT infrastructuur voldoet aan de procedures van de organisatie voor beveiliging, back-up en disaster recovery (bijvoorbeeld bij een storing of schade).
A17	dataToegang	De juiste gebruikers hebben toegang tot de juiste data om hun werkzaamheden uit te voeren.

Noot: Voor de volledige vragenlijst inclusief antwoordopties, zie bijlage C.

De ambitie is om 125 respondenten te verzamelen voor dit onderzoek. Op basis van deze steekproefomvang bestaat het vertrouwen in de generaliseerbaarheid naar de gehele populatie, alle MKB organisaties in Nederland. De 125 respondenten worden getracht te bereiken door inzet van eigen netwerk van een groep van 4 studenten, via berichten op LinkedIn en via contacten vanuit werkgevers van studenten.

Gegevensanalyse

Beschrijvende statistieken. Wanneer de gegevens zijn verzameld, worden hier allereerst beschrijvende statistieken van gegenereerd. Denk hierbij onder andere aan de sector waarin de organisatie actief is, het aantal medewerkers van de organisatie en de regio waar het hoofdkantoor gevestigd is. Deze data wordt verzameld om inzicht te krijgen in de validiteit van de steekproef.

Betrouwbaarheidsanalyse. Vervolgens wordt een betrouwbaarheidsanalyse uitgevoerd door de Cronbachs Alpha te berekenen voor de verschillende categorieën vragen. Hiermee wordt getoetst of de verschillende items (antwoorden op vragen) uit een categorie samen één schaal mogen vormen en dus daadwerkelijk samen binnen één categorie vallen. Een Cronbachs Alpha van .7 of hoger wordt geaccepteerd; mocht voor een categorie een Cronbachs Alpha van lager dan .7 gelden, worden 1 of meerdere items verwijderd waardoor de Cronbachs Alpha naar alle waarschijnlijkheid stijgt.

Normaalverdeling. De items worden gecontroleerd op normaalverdeling. Een randvoorwaarde voor het uitvoeren van de kwantitatieve analyse, is dat alle items normaal verdeeld zijn. Voor items waar dit niet het geval is, zal een transformatie toegepast worden (Field, 2009).

Data analyse. Voor de validatie van de verschillende dimensies wordt een exploratieve factoranalyse gebruikt. Met deze analyse wordt bekeken of alle onderzochte variabelen daadwerkelijk onderdeel uitmaken van de subdimensies van Tangible Resources, en hoe deze variabelen onderverdeeld kunnen worden in constructen. De factoranalyse wordt in een aantal stappen uitgevoerd. Zo wordt allereerst bekeken welke variabelen relevant zijn voor de factoranalyse door te kijken naar communalities. Vervolgens wordt bekeken hoe veel factoren gecreëerd worden. Ten slotte wordt bekeken bij welke factoren de variabelen behoren. Dit iteratieve proces zal naar alle waarschijnlijkheid meerdere keren doorlopen worden om tot een resultaat te komen waarmee de onderzoeksvraag beantwoord kan worden.

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

Mogelijke beperkingen aangaande de betrouwbaarheid van het onderzoek bestaan in de selectie van de steekproef. Deze is niet willekeurig, studenten gebruiken direct of indirect hun eigen netwerk, daardoor is het mogelijk dat er bias ontstaat in de steekproef. Zo kunnen bedrijven eerder geneigd zijn te reageren op de oproep tot onderzoek, wanneer ze zich al enigszins verdiept hebben in DA en dus meer geneigd zijn om een grotere inzet van DA te rapporteren. Zoals eerder gemeld wordt getracht de variatie te waarborgen, door 4 studenten respondenten te laten aanleveren. Doordat de 4 studenten allen een ander netwerk (andere sector en ander werkgebied) raadplegen, is er een grotere kans op een gevarieerde steekproef, wat de generaliseerbaarheid ten goede komt.

Ook de respondent binnen het MKB bedrijf die de vragenlijst invult, kan een betrouwbaarheidsrisico met zich meebrengen. Verschillende rollen die de vragenlijst invullen, kunnen leiden tot verschillende antwoorden. Daarom zal Functie meegenomen worden als covariabele. Hiermee kan bekeken worden of de ingevulde waarden significant afwijken bij vragenlijsten ingevuld door iemand uit de directie, ten opzichte van ingevulde waarden ingevuld door iemand niet werkzaam als directielid. Daarnaast is ook een controlevraag in de vragenlijst opgenomen, waarmee gevalideerd wordt dat de respondent de vragenlijst serieus heeft ingevuld.

De gebruikte vragenlijst is specifiek voor dit onderzoek opgesteld en dus nog niet eerder getoetst op betrouwbaarheid en validiteit. Daarentegen is de vragenlijst opgebouwd op basis van een theoretisch kader, opgesteld door middel van uitvoerig onderzoek. Hierdoor is de verwachting dat de vragenlijst betrouwbaar en daarmee bruikbaar voor dit onderzoek is.

Ethiek. Bij het vergaren van de data via de vragenlijst worden geen persoonlijke gegevens verzameld. Daarnaast wordt zo goed mogelijk getracht de anonimiteit te waarborgen. De data wordt minimaal 10 jaar bewaard, volgens de richtlijnen beschreven in de Nederlandse gedragscode wetenschappelijke integriteit. Daarnaast wordt expliciet om toestemming gevraagd voor deelname aan het onderzoek, aan de respondenten. Op elk moment, zonder opgave van reden, is het mogelijk voor respondenten om zich terug te trekken uit het onderzoek en hun response te verwijderen uit de database. Het onderzoek is onafhankelijk, wat voor dit onderzoek inhoudt dat het onderzoek niet geleid wordt door buiten-wetenschappelijke factoren, in dit geval organisatieagenda's.

Resultaten

In dit hoofdstuk worden de resultaten van de kwantitatieve analyse besproken. Dit hoofdstuk bestaat achtereenvolgens uit een paragraaf Data inspectie, waar ook de Missing Value Analysis en betrouwbaarheidsanalyse onderdeel van uitmaken, een paragraaf Descriptives, waar zowel de achtergrondvariabelen als de itemscores besproken worden, en een paragraaf waar de resultaten van de daadwerkelijke factoranalyse besproken worden.

Data inspectie

Er hebben in totaal 125 respondenten geparticipeerd aan dit onderzoek. Van deze responses zijn 99 valide; 26 responses (20.8%) zijn uitgesloten aangezien deze niet voldoen aan het inclusie criteria 'valt onder MKB'. Deze responses zijn ingevuld door een respondent van een bedrijf met meer dan 250 medewerkers, of bij deze response is de bedrijfsgrootte niet ingevuld.

Daarnaast waren er van de overige steekproef (N=99), 12 responses (12.1%) niet bruikbaar aangezien alleen vragen zijn beantwoord over de achtergrondvariabelen. Deze responses zijn niet meegenomen in de analyses.

Vervolgens waren er 7 respondenten (8.0%) die op 1 of meerdere vragen (1, 2 of 3) geen antwoord wilden of konden geven. Aangezien bij deze responses het percentage missing data minimaal was, is besloten om

bij deze respondenten de missende data te imputeren met een gemiddelde. Uiteindelijk zijn er 87 valide responses gebruikt voor de analyses.

Missing Value Analysis. Een Missing Value Analysis (MVA) is uitgevoerd om te bepalen of er opvallende verschillen te zien zijn tussen de respondenten die de vragenlijst volledig hebben ingevuld en respondenten die vroegtijdig zijn gestopt. Hierbij is er gekeken of de missende waardes berusten op toeval of dat er een relatie was tussen vroegtijdig stoppen en de data. De resultaten van de MVA zijn in Tabel 4 te zien.

Tabel 4

Resultaten Missing Value Analysis

		N volledig	N Missing	% Missing
Functie	Directie	74	9	11.6%
	Niet-Directie	13	1	9.5%
Kanaal	Eigen netwerk	48	4	7.7%
	Klanten werkgever	18	1	5.3%
	LinkedIn	21	5	25%
Regio	Noord	37	3	7.5%
	Midden	33	6	19.5%
	Zuid	17	1	5.6%
Bedrijfs grootte	1 werknemer	12	2	14.3%
	Micro	26	5	18.8%
	Klein	22	1	4.3%
	Middelgroot	27	3	10.0%
Sector	Industrie & Energie	14	2	15.8%
	Commercieel	71	6	8.3%
	Niet-commercieel	2	1	50%
Rechtsvorm	Eenmanszaak	18	2	10%
	Besloten vennootschap	57	7	12.3%
	Naamloze vennootschap	1	0	0%
	Vennootschap onder firma	7	1	12.5%
	Coöperatie / Onderlinge waarborgmaatschappij	2	0	0%
	Stichting	2	0	0%

Zoals te zien in Tabel 4, komt uit de MVA naar voren dat er geen aantoonbare relatie is tussen de missende waardes en de functie van de respondent. Voor interpretatiedoeleinden is functie van de respondent gegroepeerd op Directie of Niet-directie. Directieleden waren even geneigd om door te gaan (of vroegtijdig te stoppen) in vergelijking met niet-directieleden (allen $p > .05$). Vervolgens is er een relatie gezien tussen missende waardes en wijze van werven. Respondenten die via LinkedIn zijn geworven, waren eerder geneigd om vroegtijdig te stoppen: responsen van LinkedIn waren 25% onvolledig (totale N=21), responses

vanuit het eigen netwerk waren 7.7% onvolledig is (totale N=48) en responses die geworven waren via klanten van de werkgever waren 5.3% onvolledig (totale N=18). Hiernaast is gekeken naar de provincies waar de organisatie gevestigd is. Deze zijn voor interpretatiedoeleinden gegroepeerd op regio's noord, midden en zuid. Hierbij wordt een relatie gezien tussen de missende waardes en regio: respondenten uit de regio midden waren mogelijk eerder geneigd om vroegtijdig te stoppen. In regio midden was 19.5% (totale N=33) van de responses onvolledig, waar dit voor regio noord 7.5% (totale N=37) is en voor regio zuid 5.6% (totale N=17) is. Ook met betrekking tot de bedrijfsgrootte zijn relaties met missende waarden geconstateerd. Respondenten van micro organisaties (18.8%, totale N=26) en organisaties met 1 werknemer (14.3%, totale N=12) waren eerder geneigd om vroegtijdig te stoppen dan respondenten van kleine organisaties (4.3%, totale N=22) en middelgrote organisaties (10.0%, N=27).

Ten slotte is gekeken naar de sector en rechtsvorm van de organisaties van de respondenten. Hier zijn door beperkte groepsgrootten echter geen aantoonbare relaties gevonden met missende waarden.

Betrouwbaarheid. De algehele schaal voor Tangible resources is betrouwbaar, Cronbachs Alpha voor de 17 items is .79. Er is dus sprake van voldoende interne consistentie tussen de items, aangezien de waarde boven de .70 ligt (Saunders et al., 2016).

Descriptives

Achtergrondvariabelen. Beschrijvende statistieken van de achtergrondvariabelen staan in Tabel 5. Opvallend is dat het overgroot gedeelte van de respondenten is geworven vanuit het eigen netwerk (55.2%). Daarnaast was het overgroot gedeelte (85.1%) directeur (Directeur/eigenaar, CEO, CTO, CIO, CDO, CFO of COO). Vervolgens is er een goede spreiding te zien in de respondenten van verschillende bedrijfsgrootten en regio's. Het overgrote gedeelte (81.6%) van de respondenten is werkzaam voor een organisatie in de commerciële sector. De meeste respondenten zijn werkzaam voor een organisatie met rechtsvorm Besloten Vennootschap (65.5%).

Tabel 5*Descriptives achtergrondvariabelen*

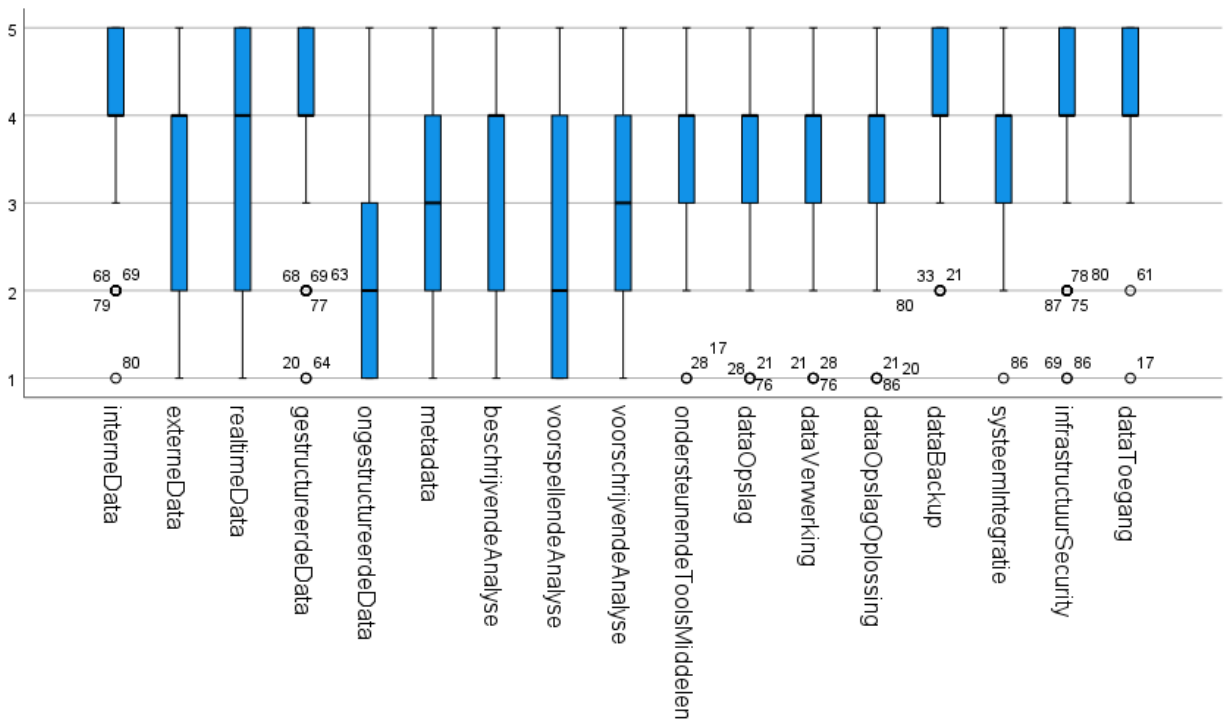
		N	%
Functie	Directie	74	85.1%
	Niet-Directie	13	14.9%
Kanaal	Eigen netwerk	48	55.2%
	Klanten werkgever	18	20.7%
	LinkedIn	21	24.1%
Regio	Noord	37	42.5%
	Midden	33	37.9%
	Zuid	17	19.5%
Bedrijfsgrootte	1 werknemer	12	13.8%
	Micro	26	29.9%
	Klein	22	25.3%
	Middelgroot	27	31.0%
Sector	Industrie & Energie	14	16.1%
	Commercieel	71	81.6%
	Niet-commercieel	2	2.3%
Rechtsvorm	Eenmanszaak	18	20.7%
	Besloten vennootschap	57	65.5%
	Naamloze vennootschap	1	1.1%
	Vennootschap onder firma	7	8.0%
	Coöperatie / Onderlinge waarborgmaatschappij	2	2.3%
	Stichting	2	2.3%

Beschrijving itemscores. Een weergave van de spreiding in de data van de itemscores is te zien in Figuur

2.

Figuur 2

Boxplotten itemscores



Er is gebruik gemaakt van een 5-punts Likert scale bij de afname van de items over Tangible resources. Dit zijn dus ordinale variabelen, echter worden deze voor analysedoeleinden als numeriek beschouwd. In Figuur 2 is te zien dat een aantal respondenten op meerdere items lager scoren dan de rest van de steekproef. Zo zijn response 21, 28, 69, 86 meerdere keren terug te zien als outlier.

Een voorwaarde van een Principal Component Analysis is dat de items normaal verdeeld zijn. Zoals te zien in Figuur 2, zijn er mogelijk meerdere items scheef verdeeld. Om normaliteit te toetsen zijn voor alle items de Z-scores van de Skewness en Kurtosis waarden berekend. Hierbij is een cut-off score van 3 gebruikt (Field, 2009): items waarbij de Z-scores buiten het bereik van -3 en 3 liggen, worden als scheef verdeeld beschouwd. Dit is het geval voor 5 variabelen (zie Figuur 2), te weten interneData, gestructureerdeData, dataBackup, infrastructuurSecurity en dataToegang.

Deze items zijn allen scheef verdeeld naar links; hogere antwoorden zijn vaker ingevuld. Aangezien er nu niet wordt voldaan aan de randvoorwaarden van PCA is besloten om deze 5 items te transformeren

door middel van een square transformation. Na deze transformatie is voor deze 5 getransformeerde items opnieuw de normaalverdeling gecontroleerd. Deze is voor de 5 items acceptabel. Een overzicht van de Z-scores van de Skewness en Kurtosis van de variabelen met itemscores is te zien in Tabel 6.

Tabel 6

Overzicht Z-scores Skewness en Kurtosis itemscores

Item	Oorspronkelijk		Getransformeerd (square)	
	Skewness	Kurtosis	Skewness	Kurtosis
interneData	-4,57	1,36	-2,77	-1,26
externeData	-1,12	-2,06		
realtimeData	-1,27	-2,52		
gestructureerdeData	-4,48	1,92	-2,16	-1,24
ongestructureerdeData	2,60	-1,10		
metadata	-0,39	-2,51		
beschrijvendeAnalyse	-2,34	-1,42		
voorspellendeAnalyse	2,05	-1,71		
voorschrijvendeAnalyse	0,02	-2,46		
ondersteunendeToolsMiddelen	-1,61	-0,47		
dataOpslag	-2,58	-0,57		
dataVerwerking	-2,74	-0,40		
dataOpslagOplossing	-2,31	-0,80		
dataBackup	-3,62	0,74	-1,97	-1,41
systeemIntegratie	-1,42	-1,17		
infrastructuurSecurity	-3,89	0,51	-2,02	-1,78
dataToegang	-4,73	6,11	-1,42	-0,42

Factoranalyse

Er is gebruik gemaakt van een exploratieve factoranalyse. Hiervoor is een Principal Component Analysis (PCA) uitgevoerd op de 17 items, waarvan 5 items getransformeerd zijn (N=87). Deze PCA is uitgevoerd door middel van een oblique rotation (direct oblimin) met $\delta = 0$, dit omdat op basis van de theorie verwacht wordt dat factoren gecorreleerd zullen zijn. De Kaiser-Meyer-Olkin (KMO) heeft bevestigd dat het sample valide is voor analyse, $KMO = .78$. Bartlett's test of sphericity levert een Chi-square van 580.253, $p < .001$, wat voldoende correlatie tussen de items aantoont.

Relevantie items. In Tabel 7 zijn de Communalities na extractie te zien van de verschillende variabelen. Variabelen met een communality $<.4$ hebben een verwaarloosbare bijdrage aan het meten van onderliggende factoren (Field, 2009). Op basis van deze grenswaarde is besloten om de variabele externeData niet mee te nemen in het vervolg van de factoranalyse.

Tabel 7

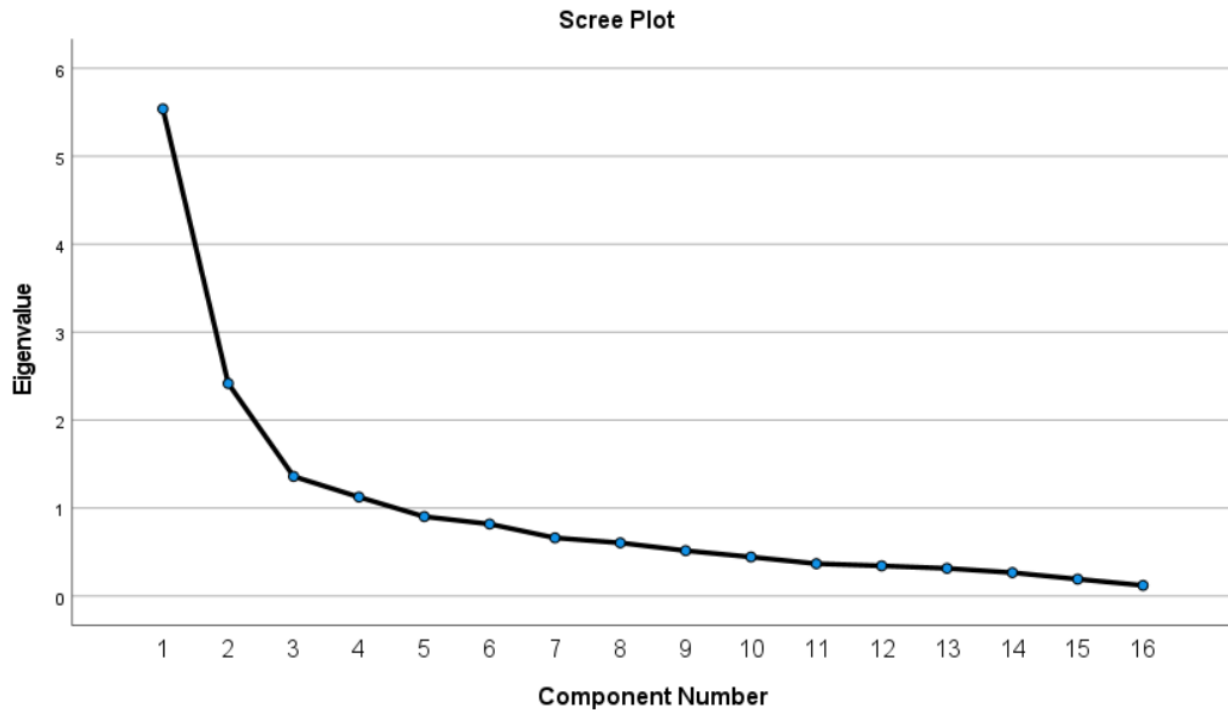
Communalities na extractie

	Extraction
interneData	.713
externeData	.383
realtimeData	.597
gestructureerdeData	.595
ongestructureerdeData	.713
metadata	.561
beschrijvendeAnalyse	.422
voorspellendeAnalyse	.725
voorschrijvendeAnalyse	.634
ondersteunendeToolsMiddelen	.608
dataOpslag	.697
dataVerwerking	.756
dataOpslagOplossing	.613
dataBackup	.731
systeemIntegratie	.573
infrastructuurSecurity	.786
dataToegang	.624

Factoren. Vier componenten hebben een eigenvalue hoger dan 1, welke samen 65.2% van de variantie verklaren. Het buigpunt in de Scree Plot, zie Figuur 3, is echter te zien bij component 3. Ook wordt door component 3 (8.5%) en component 4 (7.0%) weinig extra variantie verklaard ten opzichte van component 1 (34.6%) en component 2 (15.1%).

Figuur 3

Scree plot

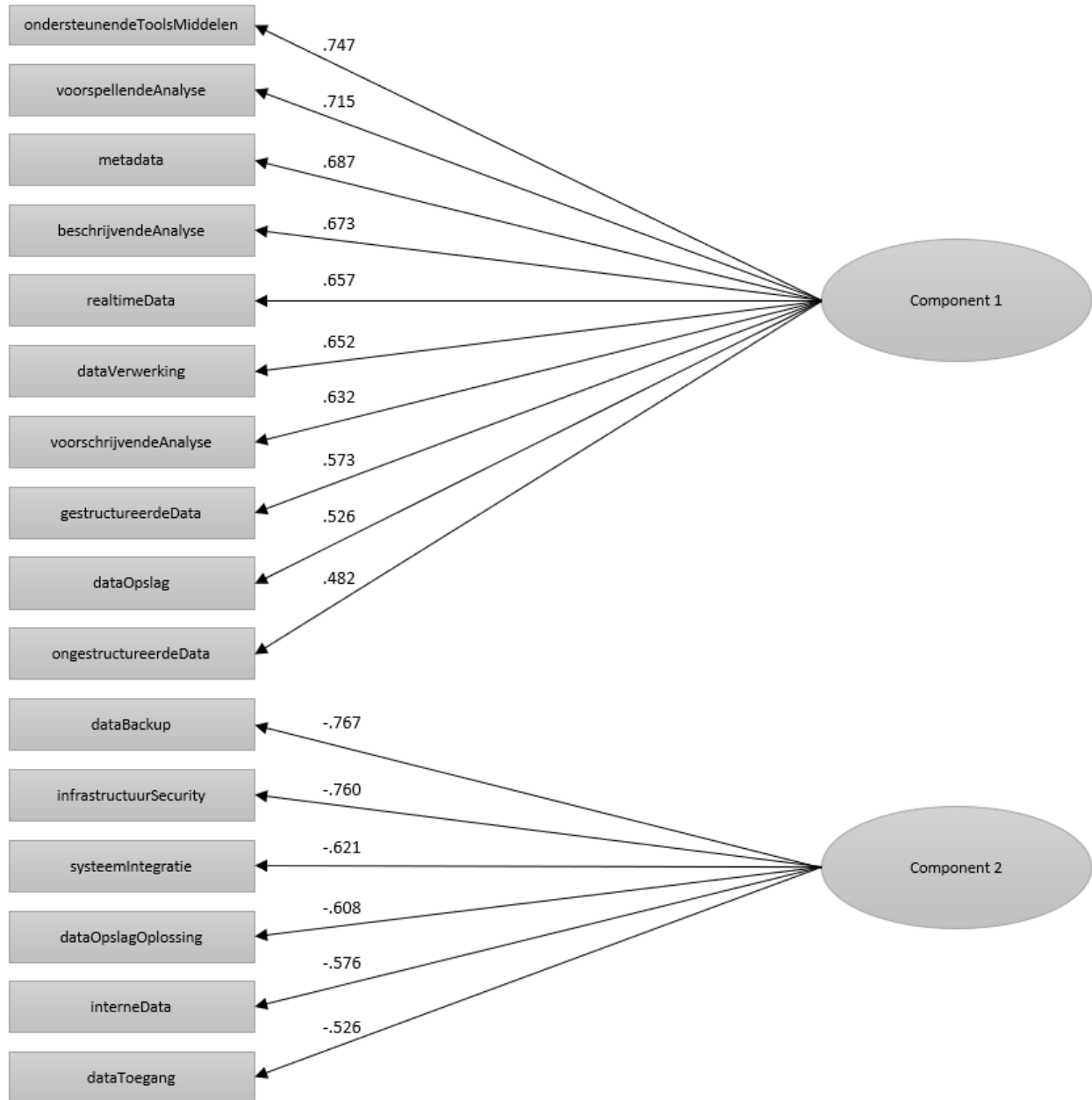


De factoranalyse is meerdere keren uitgevoerd, voor een vast aantal van 2, 3 of 4 componenten. Voor de volledigheid zijn in bijlage A de pattern matrices te zien van deze analyses. Op basis van de verklaarde variantie en het buigpunt in de Scree Plot wordt gekozen voor 2 componenten. In de discussie wordt deze keuze verder toegelicht. De cumulatieve verklaarde variantie met een vast aantal van 2 componenten is 49.7%.

Itemverdeling. In Figuur 4 is de verdeling van items over factoren te zien, bij een vast aantal van 2 factoren.

Figuur 4

Componenten en ladingen



Discussie, conclusies en aanbevelingen

Het doel van dit onderzoek was om te bepalen welke subdimensies onderdeel zijn van de dimensie Tangible Resources, waarmee een DA Capability gecreëerd kan worden, in het Nederlandse MKB. Hierbij is het resultaat dat de subdimensie Externe data onvoldoende bijdraagt aan deze opbouw. Alle overige onderzochte subdimensies maken onderdeel uit van de dimensie Tangible Resources, echter is de indeling van deze subdimensies in twee constructen afwijkend van de literatuur en het gecreëerde conceptuele model.

Discussie – reflectie

Factoren. De factoranalyse is meerdere keren uitgevoerd, voor een vast aantal van 2, 3 en 4 factoren. Resultaten van de factoranalyses met 3 en 4 factoren laten echter onlogische resultaten zien. Bij de factoranalyse met 4 factoren worden twee factoren gecreëerd met één of enkele subdimensies, wat niet mogelijk is (Field, 2009). Daarnaast worden bij de factoranalyse met 3 factoren meerdere variabelen bij meerdere factoren ingedeeld. Dit duidt op minder onderscheidende factoren. Op basis van theorie worden hier ook variabelen samengevoegd die logischerwijs niet bij elkaar horen, bijvoorbeeld Data Toegang en Voorspellende Analyse, of Data Opslag en Voorschrijvende Analyse. Mede op basis van bovengenoemde verklaring, inclusief de verklaarde variantie van factor 3 en 4 en het buigpunt in de Scree Plot, is gekozen voor 2 factoren. Deze factoranalyse laat ook logische resultaten zien.

De factoranalyse heeft dus geresulteerd in 2 factoren. Dit in tegenstelling tot het gebruikte conceptuele model waarin 4 subdimensies van Tangible Resources zijn opgenomen. Een mogelijke verklaring voor het beperkte aantal factoren ten opzichte van het conceptuele model, is dat het MKB meer gebruik maakt van standaardsoftwarepakketten voor DA. Dit omdat kennis over DA in mindere mate aanwezig is dan bij grotere organisaties en dus meer gebruikt gemaakt moet worden van de standaardpakketten aanwezig in de markt (Coleman et al., 2016). Deze standaardpakketten leveren een standaard set aan functionaliteit die in zijn geheel wel of niet gebruikt wordt. Op deze manier wordt minder

fijnmazig nagedacht over de inzet van DA binnen het MKB, waardoor het onderscheid tussen bijvoorbeeld beschrijvende analyse, voorspellende analyse en voorschrijvende analyse niet gemaakt wordt.

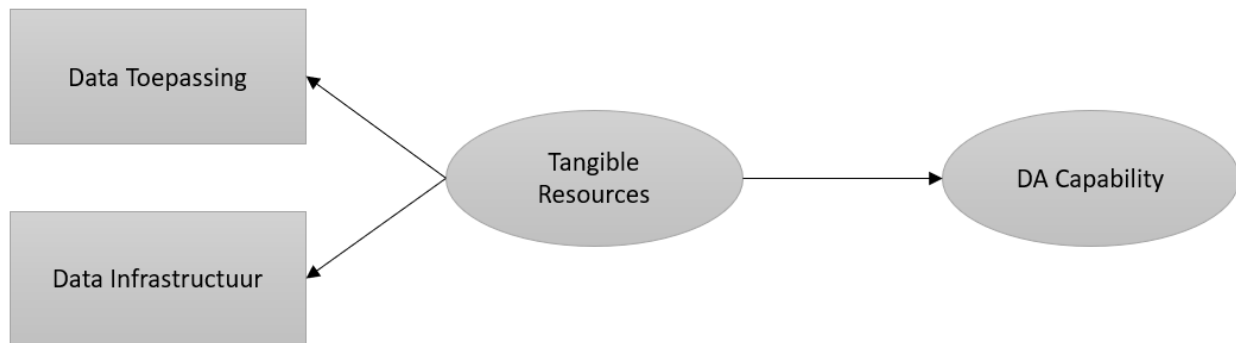
Relevantie subdimensies. De factoranalyse laat zien dat vrijwel alle onderzochte subdimensies relevant zijn voor de opbouw van een DA Capability binnen het MKB. Dit is in lijn met de verwachting die gecreëerd is op basis van de literatuur. Ondanks dat de indeling in factoren wellicht wat anders is dan bij grote bedrijven, kan het conceptuele model dus toch grotendeels gegeneraliseerd worden naar het MKB.

Op basis van de literatuur is de verwachting gecreëerd dat binnen het MKB externe data, data gecreëerd buiten de organisatie, minder gebruikt wordt dan interne data (Aydiner et al., 2019). Dit is ook terug te zien in de resultaten. De subdimensie Externe Data draagt onvoldoende bij aan de opbouw van een DA Capability voor het MKB. Een mogelijke verklaring hiervoor kan zijn dat het MKB onvoldoende expertise in huis heeft om externe data te extraheren, integreren met interne data en te gebruiken (Coleman et al., 2016). Standaardsoftwarepakketten die gebruikt worden leveren standaard analysemogelijkheden die gericht zijn op het gebruik van interne data, buiten deze standaardsoftwarepakketten is het MKB over het algemeen minder in staat om te investeren in maatwerkanalyse (Muthaih & Venkatesh, 2012).

Itemverdeling. De twee gecreëerde factoren kunnen worden geïnterpreteerd als Data Toepassing en Data Infrastructuur. Hierbij is te zien dat eerder gedefinieerde subdimensies van Data Collection en Data Analytics vallen onder Data Toepassing, waar subdimensies van Data Architecture en Technology Infrastructure onder Data Infrastructuur vallen. Data Toepassing en Data Infrastructuur zijn beiden subdimensies van Tangible Resources en dragen op die manier bij aan de opbouw van een DA Capability voor het MKB (te zien in Figuur 5).

Figuur 5

Nieuw gecreëerde subdimensies van Tangible Resources



Op bovenstaand statement is echter wel 1 uitzondering. De subdimensie Interne Data valt onder de factor Data Infrastructuur, waar deze in het conceptuele model onderdeel uitmaakt van de subdimensie Data Collection en dus volgens bovenstaande interpretatie onder Data Toepassing zou moeten vallen. Een mogelijke verklaring voor deze uitzondering is interpretatie van de vraag door de respondenten. In deze vraag wordt gesproken over het gebruik van interne databronnen in plaats van het gebruik van interne data. Deze vraag zou dus door respondenten zo opgevat kunnen zijn, dat interne systemen aanwezig zijn die interne data bevatten. Aanwezige systemen valt logischerwijs onder Data Infrastructuur, in plaats van onder Data Toepassing.

Het gebruik van ongestructureerde data is een variabele die laadt op beide factoren. Echter is de hoogste lading te zien bij de factor Data Toepassing. Ook is het op basis van de theorie logischer om deze variabele bij Data Toepassing te scharen. Andere variabelen die in het conceptueel model bij dezelfde subdimensie horen (gestructureerde data en realtime data) vallen ook onder de factor Data Toepassing.

Voorspellende analyse en voorschrijvende analyse zijn twee variabelen die naar verwachting minder zouden bijdragen aan de opbouw van een DA Capability binnen het MKB. Echter is in de resultaten te zien dat deze variabelen wel degelijk bijdragen, aan de factor Data Toepassing. Een verklaring zou het gebruik van standaardpakketten zijn, waar voorspellende analyses en voorschrijvende analyses onderdeel

uitmaken van de standaardfunctionaliteit (Coleman et al., 2016). Het MKB maakt dus wellicht meer gebruik van deze functionaliteit dan vooraf gedacht.

Transformatie variabelen. Voor de factoranalyse is een vijftal variabelen getransformeerd; de antwoorden op vragen voor deze variabelen waren niet normaal verdeeld. Overige variabelen zijn niet getransformeerd, deze waren normaal verdeeld. Om impact op de factoranalyse uit te sluiten is deze ook gedraaid met uitsluitend ongetransformeerde variabelen en uitsluitend getransformeerde variabelen. In beide gevallen is geen impact op de verdeling van items in factoren gevonden. Daarnaast wordt in beide gevallen niet aan de randvoorwaarde voldaan dat alle itemschalen normaal verdeeld zijn (Field, 2009). Dit motiveert ook de keuze om voor de factoranalyse te kiezen waarbij 5 variabelen zijn getransformeerd. Zie Bijlage B voor de pattern matrices van de factoranalyse met uitsluitend ongetransformeerde en getransformeerde variabelen.

Limitaties. Dit onderzoek kent een aantal limitaties, maar ook een aantal sterke punten. Wanneer gekeken wordt naar de limitaties dan zou er mogelijk sprake kunnen zijn van selectie bias. Door respondenten te verzamelen uit een direct of indirect eigen netwerk van de studenten of klanten van werkgevers, worden voornamelijk respondenten verzameld die geïnteresseerd zijn in DA en mogelijk ook DA dus ook meer inzetten binnen hun organisatie. Echter is in de resultaten terug te zien dat ook respondenten aan het onderzoek hebben geparticipeerd die over het algemeen lager scoren op de verschillende items, en dus een beperktere inzet van DA binnen hun organisatie tonen. Er is dus sprake van een steekproef bestaande uit respondenten die zowel een beperkte inzet in DA tonen, als respondenten die een uitgebreide inzet van DA tonen binnen hun organisatie. Hiermee is aangetoond dat er waarschijnlijk geen sprake is van selectie bias. Op basis van de beschrijvende gegevens verzameld over de respondenten valt ook aan te nemen dat de steekproef representatief is. Zo is een spreiding in bedrijfsgrootten te zien die grotendeels overeenkomt met de spreiding in bedrijfsgrootten van de gehele populatie van het MKB Nederland (Statistiek, 2021). Ditzelfde geldt voor de sectoren en rechtsvormen van de organisaties waar respondenten werkzaam zijn. Daarnaast is het zo dat de respondenten grotendeels (85.1%) onderdeel zijn

van de directie van hun organisatie. De directie van een organisatie heeft over het algemeen een breed beeld van de gehele organisatie, wat de waarheidsgetrouwheid van deze responsen aannemelijker maakt. Hiernaast kan geconcludeerd worden dat missende waarden niet willekeurig voorkomen. Respondenten geworden via kanaal LinkedIn, gevestigd in regio Midden-Nederland of werkzaam voor een micro-organisatie waren eerder geneigd om de vragenlijst niet volledig in te vullen. Dit kan leiden tot een mogelijke bias in de dataset. Ook is de steekproef beperkt. In de analyse zijn 87 valide responses gebruikt, waar de literatuur aangeeft dat een aantal van minimaal 300 respondenten aanbevolen is voor een factoranalyse. Echter, heeft de Kaiser-Meyer-Olkin (KMO) bevestigd dat de steekproef wel valide is voor analyse, $KMO = .78$. Een andere limitatie van dit onderzoek is de validiteit van de vragenlijst. Deze is voor dit onderzoek opgesteld en nog niet eerder gebruikt, wat de validiteit van de vragenlijst onzeker maakt. Wel kan geconcludeerd worden de vragenlijst betrouwbaar is (Cronbachs alpha = .79) en de vragenlijst is daarnaast gebaseerd op eerder wetenschappelijk onderzoek. In het gebruikte wetenschappelijk onderzoek om de vragenlijst te creëren, zijn vragenlijsten gebruikt waarvan validiteit is aangetoond. Ten slotte is een limitatie van dit onderzoek dat het zich richt op een beperkt onderdeel van een groter conceptueel model, waar ook de dimensies Intangible Resources, Architectuur en Strategie onderdeel van uitmaken. Daarentegen is dit onderzoek wel vernieuwend; het vult een duidelijke knowledge gap in. Bestaand onderzoek focust zich op grotere organisaties, dergelijk onderzoek is voor het MKB nog niet uitgevoerd, het raamwerk opgesteld door Moonen en collega's was nog niet getoetst.

Conclusies

Op basis van de resultaten, kan geconcludeerd worden dat het antwoord op de hoofdvraag *Welke subdimensies vormen een valide meetinstrument voor het meten van Tangible Resources voor een DA Capability binnen het MKB?* is dat alle subdimensies, meegenomen in dit onderzoek onderdeel zijn van Tangible Resources, op externe data na. Echter is de groepering van deze subdimensies wel anders dan in het conceptueel model. Dit laat dus zien dat er grote overeenkomsten zijn tussen de opbouw van een DA

Capability op basis van Tangible Resources binnen het MKB en de opbouw voor grotere organisaties. Een DA Capability wordt daarbij opgebouwd door de subdimensies Data Toepassing en Data Infrastructuur.

Aanbevelingen voor de praktijk

Voor IT leveranciers wordt allereerst aanbevolen om de resultaten van dit onderzoek te gebruiken om het Data Analytics (DA) dienstportfolio te valideren en/of uit te breiden. Daarnaast kan dit onderzoek voor een salesorganisatie gebruikt worden als gespreksinitiator richting het MKB. Wetenschappelijk onderzoek toont immers de meerwaarde van DA aan, ook voor het Nederlandse MKB. Ten slotte wordt aanbevolen om dit onderzoek te gebruiken als het startpunt voor de adoptie van DA. Het MKB heeft vaak minder financiële ruimte om te investeren in DA en dus is focus op belangrijke aspecten van DA voor het MKB benodigd.

Aanbevelingen voor verder onderzoek

Voor verder onderzoek wordt allereerst aanbevolen om ditzelfde onderzoek met een grotere groep respondenten uit te voeren. Op deze manier kan wellicht onderscheid gemaakt worden in de resultaten voor bijvoorbeeld verschillende sectoren of bedrijfsgrootten binnen het MKB. Ook wordt aanbevolen om de vragenlijst te valideren. Na dit onderzoek is nog onbekend of dat de gebruikte set aan (sub)dimensies volledig is om een DA Capability op basis van Tangible Resources te creëren. Op basis van dit onderzoek kan wel geconcludeerd worden dat de verschillende DA (sub)dimensies laden op de dimensie Tangible Resources, maar het is nog onbekend of dit alle (sub)dimensies zijn die een DA Capability op basis van Tangible Resources vormen. Ten slotte wordt aanbevolen om gericht onderzoek te doen naar de impact van de inzet van DA Capabilities op de bedrijfsperformance en hier ook de andere subdimensies Intangible Resources, Governance en Strategie in mee te nemen. Op deze manier kan de kennis op het gebied van DA in het MKB nog verder uitgebreid worden.

Literatuur

- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International journal of production economics*, 182, 113-131.
- Ashrafi, A., Ravasan, A. Z., Trkman, P., & Afshari, S. (2019). The role of business analytics capabilities in bolstering firms' agility and performance. *International Journal of Information Management*, 47, 1-15.
- Aydiner, A. S., Tatoglu, E., Bayraktar, E., Zaim, S., & Delen, D. (2019). Business analytics and firm performance: The mediating role of business process performance. *Journal of business research*, 96, 228-237.
- Baransel, A. E., & Baransel, C. (2012, 16-20 July 2012). *Architecturing Business Intelligence for SMEs*. Paper presented at the 2012 IEEE 36th Annual Computer Software and Applications Conference.
- Bharadwaj, A. S. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation. *MIS Quarterly*, 24(1), 169-196. doi:10.2307/3250983
- Centraal Bureau voor de Statistiek (2021). *Kerncijfers MKB; bedrijfsgrootte*. Retrieved from: <https://mkbstatline.cbs.nl/#/MKB/nl/dataset/48038NED/table>
- Chatterjee, S., Rana, N. P., & Dwivedi, Y. K. (2021). How does business analytics contribute to organisational performance and business value? A resource-based view. *Information technology & people (West Linn, Or.), ahead-of-print*(ahead-of-print). doi:10.1108/ITP-08-2020-0603
- Chen, L., & Nath, R. (2018). Business analytics maturity of firms: an examination of the relationships between managerial perception of IT, business analytics maturity and success. *Information systems management*, 35(1), 62-77. doi:10.1080/10580530.2017.1416948
- Chen, Y., & Lin, Z. (2021). Business Intelligence Capabilities and Firm Performance: A Study in China. *International Journal of Information Management*, 57, 102232. doi:10.1016/j.ijinfomgt.2020.102232
- Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Reis, M. S. (2016). How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International*, 32(6), 2151-2164.
- Comuzzi, M., & Patel, A. (2016). How organisations leverage Big Data: a maturity model. *Industrial Management & Data Systems*, 116(8), 1468-1492. doi:10.1108/IMDS-12-2015-0495
- Cosic, R., Shanks, G., & Maynard, S. B. (2015). A business analytics capability framework. *Australasian Journal of Information Systems*, 19.
- Field, A. (2009). *Discovering statistics using spss third edition*.
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & management*, 53(8), 1049-1064. doi:10.1016/j.im.2016.07.004
- Huang, Y., Liu, H., Li, W., Wang, Z., Hu, X., & Wang, W. (2020). Lifestyles in Amazon: Evidence from online reviews enhanced recommender system. *International Journal of Market Research*, 62(6), 689-706. doi:10.1177/1470785319844146
- Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, 16(3), 547-578.
- Moonen, N., Baijens, J., Ebrahim, M., & Helms, R. (2019). Small Business, Big Data: An Assessment Tool for (Big) Data Analytics Capabilities in SMEs. *Academy of Management Proceedings*, 2019(1), 16354. doi:10.5465/AMBPP.2019.16354abstract
- Muthaih, K., & Venkatesh, S. (2012). A study on the barriers affecting the growth of small and medium enterprises in India. *International Journal of Research in Computer Application Management*, 2(1), 77-81.
- Nederlands Comité voor Ondernemerschap (2019). *Jaarbericht de staat van het MKB 2019*.

- Okoli, C., & Schabram, K. (2009). *Protocol for a systematic literature review of research on the Wikipedia* (Vol. 9).
- Ramadan, M., Shuqqo, H., Qtaishat, L., Asmar, H., & Salah, B. (2020). Sustainable Competitive Advantage Driven by Big Data Analytics and Innovation. *Applied sciences*, 10(19), 6784. doi:10.3390/app10196784
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students* (Vol. 7).
- Shih, W., Kaufman, S., & Spinola, D. (2007). Company Background.
- Su, X., Zeng, W., Zheng, M., Jiang, X., Lin, W., & Xu, A. (2021). Big data analytics capabilities and organizational performance: the mediating effect of dual innovations. *European Journal of Innovation Management*.
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of business research*, 70, 356-365.

Bijlagen

Bijlage A: Pattern Matrices

Tabel 8

Pattern Matrix 4 componenten

	1	2	3	4
dataOpslag	.878			
dataVerwerking	.877			
gestructureerdeData	.740			
realtimeData	.615			
ondersteunendeToolsMiddelen	.613			
interneData	.589			-.519
metadata	.539			
beschrijvendeAnalyse				
infrastructuurSecurity		-.904		
dataBackup		-.887		
dataOpslagOplossing		-.695		
systeemIntegratie	.421	-.488		
voorspellendeAnalyse			.775	
dataToegang			-.639	
voorschrijvendeAnalyse			.631	
ongestructureerdeData				.816

Tabel 9

Pattern Matrix 3 componenten

	1	2	3
dataVerwerking	.745		
gestructureerdeData	.725		
metadata	.725		
realtimeData	.691		
ondersteunendeToolsMiddelen	.686		
dataOpslag	.685		-.441
systeemIntegratie	.684		
dataOpslagOplossing	.652		
interneData	.641		
beschrijvendeAnalyse	.567		
voorspellendeAnalyse		.638	
dataBackup	.442	-.618	.405
infrastructuurSecurity	.507	-.574	.424
ongestructureerdeData		.556	
dataToegang		-.473	
voorschrijvendeAnalyse	.465	.415	.532

Tabel 10

Pattern Matrix 2 componenten

	1	2
ondersteunendeToolsMiddelen	.747	
voorspellendeAnalyse	.715	
metadata	.687	
beschrijvendeAnalyse	.673	

realtimeData	.657	
dataVerwerking	.652	
voorschrijvendeAnalyse	.632	
gestructureerdeData	.573	
dataOpslag	.526	
ongestructureerdeData	.482	.414
dataBackup		-.767
infrastructuurSecurity		-.760
systeemIntegratie		-.621
dataOpslagOplossing		-.608
interneData		-.576
dataToegang		-.526

Bijlage B: Pattern Matrices Getransformeerd/Ongetransformeerd

Tabel 11

Pattern Matrix alle variabelen getransformeerd

	1	2
ondersteunendeToolsMiddelen	.757	
dataVerwerking	.697	
beschrijvendeAnalyse	.693	
metadata	.661	
realtimeDat	.654	
voorspellendeAnalyse	.645	
voorschrijvendeAnalyse	.605	
gestructureerdeData	.603	
dataOpslag	.563	
ongestructureerdeData	.511	
infrastructuurSecurity		-.790
dataBackup		-.787
dataOpslagOplossing		-.616
systeemIntegratie		-.609
interneData		-.544
dataToegang		-.516

Tabel 12

Pattern Matrix alle variabelen ongetransformeerd

	1	2
ondersteunendeToolsMiddelen	.744	
voorspellendeAnalyse	.718	
metadata	.685	
beschrijvendeAnalyse	.679	
realtimeData	.657	
dataVerwerking	.643	
voorschrijvendeAnalyse	.635	
gestructureerdeData	.583	
dataOpslag	.514	
ongestructureerdeData	.491	.429
dataBackup		-.770
infrastructuurSecurity		-.767
systeemIntegratie		-.629
dataOpslagOplossing		-.627
interneData		-.590
dataToegang		-.442

Bijlage C: Volledige Vragenlijst Inclusief Antwoordopties

Data Analytics Capabilities in het MKB

Hartelijk dank voor uw tijd en interesse om deel te nemen aan dit onderzoek!

Met uw deelname draagt u bij aan het creëren van waardevolle inzichten voor uw organisatie, uw sector en het MKB. Ter introductie vragen wij u daarom onderstaande zorgvuldig door te lezen, zodat u weet wat u kunt verwachten.

Achtergrond

Door de toenemende globalisering en snelle technologische ontwikkelingen wordt het MKB geconfronteerd met grote uitdagingen. Waar Data Analytics veelbelovend is en potentie biedt, zijn het voornamelijk grote organisaties die hierin voortuitgang laten zien en dreigt het MKB achterop te raken met de adoptie van Data Analytics.

Dit terwijl Data Analytics organisaties in staat stelt om waarde te creëren uit data, en dit in potentie leidt tot een hogere productiviteit, winstgevendheid en concurrentievoordeel. Het MKB is en blijft echter een essentieel onderdeel van onze economie en het is voor MKB bedrijven belangrijk om mee te gaan in de digitale transformatie.

Dit onderzoek richt zich daarom op het MKB, waarbij de focus ligt op de middelen die benodigd zijn voor het ontwikkelen van de zogeheten Data Analytics Capability. Het doel hiervan is om een raamwerk te ontwikkelen dat MKB bedrijven ondersteuning biedt bij het investeren en ontwikkelen in hun eigen Data Analytics Capability, om zo ook in de toekomst concurrerend te blijven.

Het invullen van deze vragenlijst zal ongeveer 10 minuten duren.

Disclaimer

Uw ingevulde antwoorden zullen uitsluitend worden gebruikt voor dit onderzoek, dat wordt uitgevoerd in het kader van het afstudeertraject MSc. Business Process Management and IT (BPMIT) aan de Open Universiteit. Daarnaast is deze vragenlijst vertrouwelijk, volledig anoniem en bevat het verslag van uw antwoorden op de vragenlijst geen identificeerbare informatie over u.

Er zijn 20 vragen in deze enquête.

Wat is uw functie? *

Kies één van de volgende antwoorden

Kies één van de volgende mogelijkheden:

- Directeur/eigenaar
- Chief Executive Officer (CEO)
- Chief Technology Officer (CTO)
- Chief Information Officer (CIO)
- Chief Data Officer (CDO)
- Overige

Wat is de grootte van uw bedrijf? *

Kies één van de volgende antwoorden

Kies één van de volgende mogelijkheden:

- 1 werknemer
- Micro (minder dan 10 werknemers en een jaaromzet van hoogstens 2 miljoen euro of een jaarlijks balanstotaal kleiner of gelijk aan 2 miljoen euro)
- Klein (minder dan 50 werknemers, een netto-jaaromzet van hoogstens 10 miljoen euro of een balanstotaal kleiner of gelijk aan 10 miljoen euro)
- Middelgroot (minder dan 250 werknemers, een jaaromzet van hoogstens 50 miljoen euro of een jaarlijks balanstotaal kleiner of gelijk aan 43 miljoen euro)
- Groot (250+ medewerkers)

In welke sector is uw bedrijf actief? *

Kies één van de volgende antwoorden

Kies één van de volgende mogelijkheden:

- Landbouw en visserij
- Industrie
- Bouwnijverheid
- Groothandel
- Detailhandel
- Horeca
- Vervoer
- Financiën
- Adviesdiensten
- Facilitaire diensten

- Persoonlijke diensten
- Zakelijke dienstverlening
- Zakelijk beheer
- Onderwijs
- Cultuur
- Delfstoffen
- Evenementen
- Logistiek
- Marketing en/of communicatie
- Media
- Overheid
- Sport- en recreatie
- Zorg
- Overige

In welke regio is uw hoofd(kantoor) gevestigd? *

Kies één van de volgende antwoorden

Kies één van de volgende mogelijkheden:

- Drenthe (DR)
- Flevoland (FL)
- Fryslân (FR)
- Gelderland (GD)
- Groningen (GR)
- Limburg (LB)
- Noord-Brabant (NB)
- Noord-Holland (NH)
- Overijssel (OV)
- Utrecht (UT)
- Zeeland (ZL)
- Zuid-Holland (ZH)
- Overige

Wat is de rechtsvorm van uw bedrijf? *

Kies één van de volgende antwoorden

Kies één van de volgende mogelijkheden:

- Zorg
- Eenmanszaak
- Besloten vennootschap
- Naamloze vennootschap
- Maatschap

- Vennootschap onder firma
- Commanditaire vennootschap
- Vereniging met volledige rechtsbevoegdheid
- Vereniging met beperkte rechtsbevoegdheid
- Coöperatie / Onderlinge waarborgmaatschappij
- Stichting
- Overige

Data *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
----------------------------	---------------	-----------------	-------------	--------------------------	--------------------------

**De organisatie maakt gebruik
van interne databronnen voor
data-analyse.**

**De organisatie maakt gebruik
van externe databronnen voor
data-analyse.**

**De organisatie maakt gebruik
van real-time data voor data-
analyse.**

**De organisatie maakt gebruik
van gestructureerde data
(data in tabelvorm) voor data-
analyse.**

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

De organisatie maakt gebruik van ongestructureerde data (zoals video's, foto's of geluid) voor data-analyse.

De organisatie legt metadata (beschrijvende data over andere data, zoals 'Wanneer is de data vastgelegd' en 'Wie heeft de data ingevoerd') vast en gebruikt deze voor analysedoeleinden.

Data Analytics *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

De organisatie maakt gebruik van beschrijvende data-analyse om gebeurtenissen uit het verleden te begrijpen

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

De organisatie slaat data op om deze verder te verwerken en vervolgens te analyseren.

De organisatie verwerkt data om deze vervolgens te analyseren

Technologische infrastructuur *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

De organisatie heeft een data opslag oplossing, die aansluit bij de huidige wensen en eisen van de organisatie met betrekking tot data-analyse.

De organisatie heeft mogelijkheden tot back-up van data goed geregeld.

De organisatie heeft systeemintegratie zo

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
--------------------	--------	----------	------	------------------	------------------

ingeregeld dat deze geschikt
zijn voor data-analyse.

De IT infrastructuur voldoet
aan de procedures van de
organisatie voor beveiliging,
back-up en disaster recovery
(bijvoorbeeld bij een storing
of schade).

De juiste gebruikers hebben
toegang tot de juiste data om
hun werkzaamheden uit te
voeren.

Overige middelen *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
--------------------	--------	----------	------	------------------	------------------

De organisatie kan voldoende
tijd steken in Data Analytics

De organisatie heeft
voldoende financiële middelen

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

**voor het ontwikkelen van
Data Analytics producten**

Onderstaande stellingen hebben betrekking op de huidige bedrijfscultuur:

*

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

**Binnen de organisatie worden
alle besluiten gebaseerd op
resultaten van data analyse in
plaats van intuïtie en
ervaringen, dus op basis van
cijfers in plaats van
onderbuikgevoel.**

**De eigenaar of het
management van de
organisatie ondersteunt data
analytics initiatieven door
middelen beschikbaar te
stellen.**

Helemaal	Oneens	Neutraal	Eens	Helemaal	Geen
oneens				eens	antwoord

De werknemers zijn bereid om een verandering van de bedrijfscultuur door te gaan om meer datagedreven te worden.

De eigenaar of het management van de organisatie heeft vertrouwen in de vaardigheden van werknemers die data analyses uitvoeren.

De organisatie heeft een visie hoe ze kan veranderen om meer datagedreven te worden.

Onderstaande stellingen hebben betrekking op human resources waar kennis, vaardigheden en competenties onderdeel van zijn: *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal
oneens Oneens Neutraal Eens Helemaal
eens antwoord

**De organisatie beschikt intern
of extern over technische
vaardigheden en kennis voor
data analyse (zoals statistiek,
data management, reporting,
visualisatie).**

**De organisatie beschikt intern
of extern over business skills
en kennis voor data analyse
(zoals verkoop, financiën,
marketing, productie- en
bedrijfssystemen).**

**De organisatie beschikt intern
of extern over management
vaardigheden voor data
analyse (zoals het vaststellen
van KPI's (bijv. omzet, aantal
verkochte producten), het
stellen van doelen, acties
ondernemen zodat doelen
bereikt worden).**

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

De organisatie beschikt intern of extern over innovatie vaardigheden en kennis om met data analyse buiten haar comfortzone te treden (zoals risico nemen, trial and error, experimenteren).

Medewerkers beschikken in de interne en externe samenwerking over communicatievaardigheden om behoeften van de organisatie te vertalen naar concrete data analytics actiepunten en/of doelstellingen.

De organisatie biedt werknemers trainingen aan ter ontwikkeling van de eerdergenoemde kennis en vaardigheden voor data analyse (technisch, business,

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
----------------------------	---------------	-----------------	-------------	--------------------------	--------------------------

**management, innovatie,
communicatie).**

Analytics Governance *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
----------------------------	---------------	-----------------	-------------	--------------------------	--------------------------

**De organisatie houdt rekening
met de ethische aspecten van
data analyse (bijvoorbeeld
privacy, vertrouwelijkheid,
transparantie).**

**De organisatie controleert of
haar data analyse activiteiten
de wet naleven (bijvoorbeeld
de AVG, privacy
bescherming).**

**De organisatie beschikt over
standaarden en structuren die
ondersteuning bieden aan de
ethische en compliance**

Helemaal
oneens Oneens Neutraal Eens Helemaal
eens antwoord

aspecten omtrent data analyse
(bijvoorbeeld procedures,
richtlijnen, IT).

Het is duidelijk wie binnen de
organisatie bevoegd is tot en
verantwoordelijk is voor data
analyse.

De organisatie beschikt over
een structuur voor data
governance, die gericht is op
de beslissingsbevoegdheden
omtrent data (focus op
immateriële aspecten zoals;
toewijzing van iemand die
eigenaar van de data is,
iemand die de data
controleert, individuen
hebben controle over hun
eigen data).

De organisatie heeft externe
samenwerkingsverbanden

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

voor data analyse waarin ze haar middelen (bijv. werknemers, data, software) deelt met andere organisaties.

IT/Data Governance *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

De organisatie heeft geschikte procedures voor het beheren van datakwaliteit op het gebied van nauwkeurigheid (data zit dichtbij de werkelijkheid).

De organisatie heeft geschikte procedures voor het beheren van datakwaliteit op het gebied van actualiteit (data is op tijd aanwezig).

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens		eens	antwoord

De organisatie heeft een data analyse strategie.

De data analyse strategie is afgestemd op de algehele bedrijfsstrategie.

De doelstellingen van data analyse initiatieven zijn afgestemd op de algehele bedrijfsstrategie.

Deze vraag dient ter controle dat de vragenlijst is ingevuld. Hier graag het antwoord 'Vragenlijst ingevuld' selecteren.

*

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal				Vragenlijst	Helemaal
oneens	Oneens	Neutraal		ingevuld	eens

Controlevraag

Outsourcing *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
--------------------	--------	----------	------	------------------	------------------

**De organisatie besteedt niet
essentiele IT- middelen uit om
innovatie en gebruik van data
analyse te stimuleren.**

**De organisatie besteedt (iets
van) haar data analyse
activiteiten uit (bijv.
ontwikkeling van data analyse
strategie, implementatie van
data analyse, personeel dat
data analyse uitvoert, de
benodigde hardware/software
voor data analyse).**

Investerings *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal oneens	Oneens	Neutraal	Eens	Helemaal eens	Geen antwoord
--------------------	--------	----------	------	------------------	------------------

**Het management van de
organisatie is bereid te
investeren in data analyse**

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

**initiatieven die het bedrijf
verder helpen.**

**Onderstaande stellingen hebben betrekking op de operationele prestaties in de afgelopen 3 jaar,
ten opzichte van de sector: ***

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

**De productiviteit van de
organisatie was de afgelopen 3
jaar hoger dan de gemiddelde
productiviteit in de sector.**

**Het winstpercentage van de
organisatie was de afgelopen 3
jaar hoger dan het
gemiddelde winstpercentage
in de sector.**

**Het rendement op
investering van de organisatie
was de afgelopen 3 jaar hoger**

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

dan het gemiddelde in de sector.

De verkoopinkomsten van de organisatie waren de afgelopen 3 jaar hoger dan de gemiddelde verkoopinkomsten in de sector.

Onderstaande stellingen hebben betrekking op de marktprestaties in de afgelopen 3 jaar, ten opzichte van de sector: *

Kies het toepasselijke antwoord voor elk onderdeel:

Helemaal					Helemaal	Geen
oneens	Oneens	Neutraal	Eens	eens	antwoord	

De organisatie trad de afgelopen 3 jaar makkelijker nieuwe markten binnen dan het gemiddelde in de sector.

De organisatie introduceerde de afgelopen 3 jaar

Met vriendelijke groet,
Data analytics onderzoeksteam