

MASTER'S THESIS

Wat zijn de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland?

Auteur:
Marc Houtkooper

Begeleider:
Marcel de Wit

Datum:
30 april 2020

Hogeschool Utrecht / University of Applied Sciences
Postbus 182
3500 AD UTRECHT
Nederland

"The most meaningful way to differentiate your company from your competition is to do an outstanding job with information. How you gather, manage, and use information will determine whether you will win or lose"

Bill Gates, Business @ the Speed of Thought: Succeeding in the Digital Economy (1999)

Voorwoord

Voor u ligt de uitkomst van een kwalitatief onderzoek naar de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland.

Deze thesis is geschreven ter afronding van de opleiding Master of Informatics (MOI) aan de Hogeschool Utrecht (HU).

Het schrijven van dit voorwoord betekent voor mij de afronding van een lange en vooral leerzame periode. Na jaren praktijkervaring in de business intelligence en data warehousing, kwam ik ruim 3 jaar geleden bij de HU terecht voor meer theoretische verdieping. De MOI-module Business Analytics (BA) was mijn doel, maar op de één of andere manier ben ik gebleven en heb ik 6 modules doorlopen en een flink aantal skills geleerd. De voorbereiding en het schrijven van deze thesis waren daarbij het meest intensief, maar ook zeer interessant. Waarom is data warehousing zo interessant? Het antwoord komt van goeroe Ralph Kimball: Het ontwerpen van een data warehouse is, anders dan een IT-taak, een combinatie van computertechnologie, cognitieve psychologie, business inhoud en politiek!

Mijn dank gaat uit naar mijn eerste docent én begeleider: Marcel de Wit. Mede dankzij hem ben ik na BA gebleven bij de HU. En Marcel heeft mij uitstekend begeleid en gemotiveerd het afgelopen jaar tijdens mijn thesis proposal en thesis afronding. Verder gaat mijn dank uit naar mijn netwerk van collega-zelfstandigen en oud-collega's bij de organisaties waar ik gewerkt heb. Zonder hun input, had ik niet voldoende data gehad om dit onderzoek te kunnen uitvoeren. Tot slot wil ik mijn vrouw en kinderen bedanken voor alle support de afgelopen periode en de vele zondagen die ze mij hebben moeten missen.

Veel leesplezier!

Samenvatting

Al sinds midden jaren 90 hebben veel grote en middelgrote organisaties in Nederland tijd en geld gestoken in een data warehouse (EDW, DW, DWH). In die periode werden de eerste implementaties gedaan bij grote financiële dienstverleners in Nederland. Deze implementaties waren niet eenvoudig en leverden vaak niet het gewenste resultaat vanuit de business gezien. De grootste problemen rond complexiteit, doorlooptijd en kosten zijn onderzocht en beschreven in diverse publicaties aan het begin van deze eeuw. De gebrekkige samenwerking (alignment) tussen IT en de business, en daarmee het gebruik en de adoptie van een data warehouse, worden vaak vermeld als oorzaak.

Dit onderzoek richt zich niet op implementaties van data warehouses en de daarbij behorende projecten. Het focust op de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland.

Door literatuurstudie is een samengesteld model ontwikkeld voor data warehouse succesfactoren. Dit model is getoetst bij grote Nederlandse banken en verzekeringsmaatschappijen door middel van een online enquête en interviews.

De uitkomst van het onderzoek is dat de directe succesfactoren systeemkwaliteit en datakwaliteit prima op orde zijn bij de onderzochte organisaties. Daarmee zijn ook de waargenomen voordelen van een data warehouse, zoals het ondersteunen van bedrijfsprocessen en besluitvorming, van toepassing. En dit leidt tot een positieve bijdrage aan de strategische doelstellingen van deze organisaties.

De indirecte organisatorische, technische en economische factoren zijn ook grotendeels van toepassing bij de onderzochte organisaties.

Uit dit onderzoek is ook gebleken dat de beheerfactor van een data warehouse nog vaak te wensen overlaat. Onderwerpen van zorg zijn: standaardisatie en documentatie van een data warehouse, maar ook een consistente datakwaliteit vanuit de aanleverende bronsystemen.

Data Management, het beheer van data in de hele organisatie, wordt onderkend als belangrijk bedrijfsproces dat een oplossing kan bieden aan bovenstaande zaken. In de praktijk wordt er nog (te) weinig geregeld op dat vlak.

Een andere bevinding is het ontbreken van een gedegen enterprise data-architectuur (onderdeel van data management). Alle onderzochte organisaties hebben meerdere data warehouses operationeel voor meerdere bedrijfsonderdelen en -processen, wat leidt tot een versnipperd datalandschap en pluriforme rapportages. Ook zijn alle organisaties naast het data warehouse bezig met Data Science, vaak ondersteund door een Data Lake. Ook dit draagt bij aan het decentrale karakter. De intentie tot samenwerking is bij alle interviews ter sprake gekomen, maar door het ontbreken van Data Governance (onderdeel van Data Management), bijvoorbeeld in de vorm van centraal ingericht informatiemanagement en een CDO (Chief Data Officer), is het er nog niet van gekomen.

Om alle decentrale oplossingen zoveel mogelijk in lijn te brengen, verdient het dus de aanbeveling voor grote financiële dienstverleners in Nederland, om zo snel mogelijk te beginnen met (Enterprise) Data Management. Door te investeren in: Datakwaliteit, data-architectuur, metadata, reference- en masterdata, data modelering én Data Governance, krijgt men een positief effect op alle succesfactoren van een data warehouse, maar in het bijzonder de: Beheerfactoren.

Inhoudsopgave

Voorwoord	3
Samenvatting	4
Inhoudsopgave	5
1 Inleiding	7
1.1 Achtergrond	7
1.2 Probleemomschrijving	8
1.3 Onderzoeksvraag	10
1.4 Afbakening en scope	11
1.5 Aanpak	12
1.6 Contributie en toegevoegde waarde	12
2 Literatuurstudie	14
2.1 Inleiding	14
2.2 Bestaande data warehouse modellen	14
2.2.1 Empirisch model data warehouse succesfactoren	14
2.2.2 Model data warehouse succesfactoren	16
2.2.3 Model data warehouse adoptie	19
2.2.4 Samengesteld model succesfactoren	21
2.3 Recente operationele data warehouse succesfactoren	26
2.3.1 Bevestiging bestaande factoren	26
2.4 Nieuwe meetwaarden	27
2.4.1 Agile development	27
2.4.2 Data Management	27
2.4.3 Beheer	28
2.5 Conclusie Literatuurstudie	29
3 Case Study	31
3.1 Onderzoeksmethode en -ontwerp	31
3.2 Onderzoekspopulatie	32
3.3 Enquête en interview	32
3.3.1 Participanten	32
3.3.2 Vragenlijst en enquête	33
3.3.3 Interviews	33
3.3.4 Analyse	33
3.4 Ethische verantwoording	34
4 Resultaten	35
4.1 Respondenten	35
4.2 Resultaat per succesfactor	36
4.2.1 Waargenomen voordelen	36
4.2.2 Datakwaliteit	37
4.2.3 Systeemkwaliteit	38
4.2.4 Organisatorische factoren	39
4.2.5 Technische factoren	40
4.2.6 Economische factoren	41
4.2.7 Beheerfactoren	42
4.3 Ontwikkelingen	43
4.3.1 Governance	44
4.3.2 Architectuur	45
4.3.3 Decentraal	46
4.3.4 Life Cycle Management, Rationalisatie en Legacy	46
4.3.5 Data Science en Data Lake	46
4.3.6 Overige ontwikkelingen	47
5 Conclusies	49
5.1 Deelvragen en hoofdvraag	49

5.1.1	Deelvragen 1a, 1b en 1c.....	49
5.1.2	Deelvragen 2a, 2b en 2c.....	51
5.1.3	Deelvragen 3a en 3b.....	51
5.2	Onderzoeksvraag.....	52
5.3	Probleemomschrijving.....	53
6	Discussie en aanbevelingen.....	54
6.1	Afbakening en beperkingen	54
6.2	Vervolgonderzoek	55
6.3	Aanbevelingen voor de praktijk.....	55
7	Referenties	57
8	Bijlage	60
8.1	Criteria en zoektermen literatuuronderzoek.....	60
8.2	Publicaties op basis van zoekcriteria	60
8.3	Relevantietabel	64
8.4	Vragen enquête en interviews	72
8.5	Resultaten enquête	73
8.6	Data Warehouse Processen (DWP) en de Key Process Areas (KPA's).....	74

1 Inleiding

In dit inleidende hoofdstuk wordt gekeken naar de totstandkoming van de onderzoeksvraag in paragraaf 1.3 vanuit de probleemomschrijving in paragraaf 1.2 en een context in de vorm van achtergrondinformatie in paragraaf 1.1. Na de onderzoeksvraag volgen de afbakening en scope in paragraaf 1.4 en de aanpak om de vraag te kunnen beantwoorden in paragraaf 1.5. Tot slot wordt de toegevoegde waarde van dit onderzoek in paragraaf 1.6 benoemd.

1.1 Achtergrond

Al sinds midden jaren 90 hebben veel grote en middelgrote organisaties in Nederland en daarbuiten tijd en geld gestoken in een data warehouse (DW, DWH, Enterprise data warehouse, EDW). In die periode werden de eerste implementaties gedaan in de Verenigde Staten en werden de eerste onderzoeken hierover gepubliceerd. Met het toenemende succes van Enterprise Resource Planning (ERP) en Customer Relationship Management (CRM) kwam ook de behoefte om de data uit het inmiddels complexe applicatielandschap te integreren (Ramamurthy, 2007; Watson, 2004)

Succes

Data warehouses werden een succes in de westerse wereld, zoals ook blijkt uit de kaart in figuur 1.1 uit Google Trends. Hoe donkerder de kleur blauw, hoe vaker gezocht werd op de term "data warehouse".



Figuur 1.1 Aantal hits op data warehouse in de periode 2004 - 2019. Bron: Google Trends

Al aan het eind van de jaren 90 hadden de meeste grote organisaties een data warehouse of waren ze er één aan het bouwen (Curtis, 1999). Zoals Sen et al. opmerkte (2012), was er een enorme groei in het aantal data warehouse implementaties in de periode 2000-2012. Onder IT-leidinggevendens stond het in die periode als hoogste prioriteit op de IT-agenda. Ook Gartner (Swan, 2011) adviseerde CIO's om data warehousing als een noodzakelijke component te zien voor Business Intelligence (BI) en Business Analytics (BA).

Definitie

De definitie van een data warehouse evolueert in de tijd.

Bill Inmon (2005, p. 29) wordt algemeen gezien als de grondlegger van het data warehouse en omschreef het data warehouse in veel van zijn boeken en publicaties als volgt:

“A subject-oriented, integrated, non-volatile and time-variant collection of data in support of management's decisions”.

Inmon legt met deze oer-definitie de nadruk op de kenmerken van de dataverzameling:

- Onderwerp-gerelateerd: Elke afdeling en bedrijfsproces vindt er zijn/haar data, gegroepeerd per onderwerp;
- Geïntegreerd: Van alle systemen in een organisatie wordt één versie van de waarheid getoond in één (canoniek) datamodel;
- Niet volatiel: Feiten worden vastgelegd om nooit meer te veranderen;
- Tijd-geordend: Historische data worden bewaard en in de tijd vastgelegd.

Het doel van deze dataverzameling: het ondersteunen van besluitvorming.

In 2001 waren het Wixom en Watson (2001, p. 18) die de definitie wat ruimer en concreter wisten neer te zetten, met meer nadruk op de data-verbeterprocessen:

“A data warehouse (or smaller-scale data mart) is a specially prepared repository of data created to support decision making. Data are extracted from source systems, cleaned/scrubbed, transformed, and placed in data stores (Gray and Watson 1998). A data warehouse has data suppliers who are responsible for delivering data to the ultimate end users of the warehouse, such as analysts, operational personnel, and managers. The data suppliers make data available to end users either through SQL queries or custom-built decision-support applications”.

In 2007 bevestigde Ramamurthy en Sen (2007, p. 818) deze bovenstaande definities, maar gaf ook het verschil aan met andere systemen binnen een organisatie:

“Unlike on-line transaction processing (OLTP) database systems, which are detailed, application-oriented, and used to run the day-to-day operations of a firm, data warehouses are organized around subjects (e.g., customer, product, order, supplier, etc.) and store historical and summarized data for facilitating business decision making”.

Weer veel later in 2013 was het eigenlijk al niet meer noodzakelijk om een data warehouse definitie op te stellen, omdat het begrip algemeen bekend werd verondersteld. De artikelen gingen toen meer over de toepassing en functies van een data warehouse. Morris en Kimpel (2013, p. 376) schreven:

“Executives must have the right data to make strategic, tactical, and operational decisions. Data warehouses, when implemented correctly, promise to meet this need by integrating and organizing key operational data in a form that is consistent, reliable, timely, and readily available, wherever and whenever needed”.

Ondanks meerdere definities van een data warehouse, is het primaire doel van een data warehouse altijd gelijk gebleven: het ondersteunen van besluitvorming in een organisatie.

1.2 Probleemomschrijving

Sinds de theorie en de noodzaak van een data warehouse bekend werd, zijn de eerste implementatieprojecten gestart in de jaren 90. Inmiddels kan gesteld worden dat er anno nu veel praktijkervaring is opgedaan in Data Warehouse implementaties en -systemen in Nederland. Vanuit die praktijk zijn er steeds meer kritische geluiden te horen. Dit blijkt uit een groot aantal publicaties (commercieel en wetenschappelijk) met een kritische connotatie. De volgende zaken vallen hierbij op:

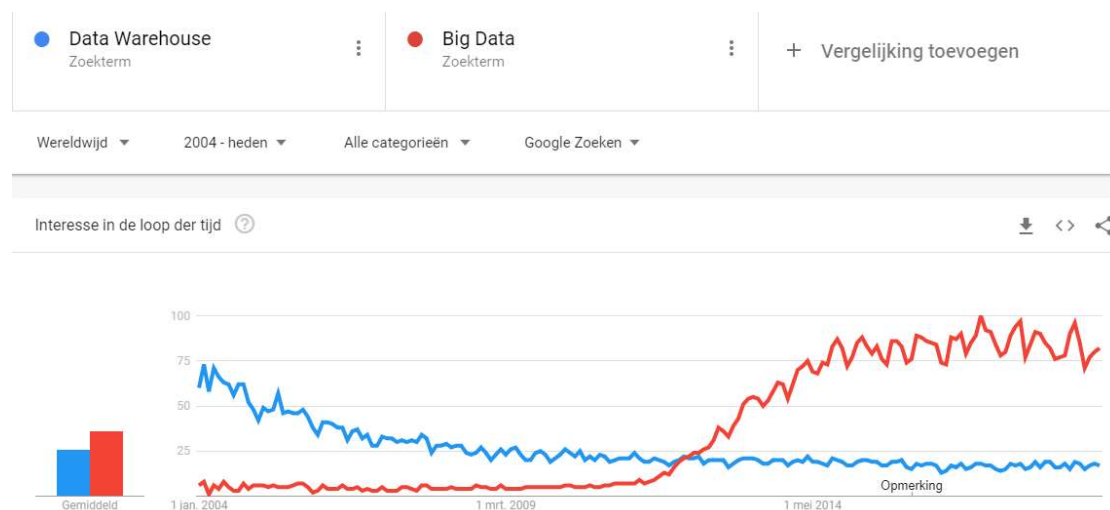
Vertrouwen en adoptie vanuit de business: Een data warehouse is meestal een IT-aangelegenheid geweest. IT was verantwoordelijk voor de laadprocessen en het (wekelijks, dagelijks) vrijgeven van de geïntegreerde data. Het vertrouwen vanuit de business is laag, want het data warehouse is (te) vaak niet juist, volledig en tijdig geladen om te dienen als basis voor de rapportages en analyses. Ook Gartner (Goodwin, 2011) merkte op dat 70 tot 80% van de business intelligence projecten mislukten, als er gekeken wordt naar de verwachtingen ten aanzien van het eindresultaat. De voornaamste reden hiervoor is gebrekkige communicatie tussen Business en IT. Maar ook omdat technenuten te weinig (juiste) vragen stellen tijdens het ontwikkeltraject.

Gevolg is dat de business vaak terugvalt op spreadsheets en Shadow IT (Nelson, 2013) wat betekent dat afdelingen zelf data gaan verzamelen met eigen tools (vaak Excel) om

rapportages te maken. Consequentie is een versnipperd datalandschap, pluriforme rapportage en meerdere versies van de waarheid.

Complexiteit & Kosten: In de basistheorieën van de 90-er jaren goeroes: Bill Inmon (2005), Ralph Kimball (1996) en later Dan Linstedt (2015) staat de “one version of the truth” centraal. De gedachte om alle systeemdata binnen een organisatie centraal en geïntegreerd op te slaan blijkt in de praktijk een complexe uitdaging. Dit resulteerde in grote en lange implementatieprojecten die vaak mislukken (Ramamurthy, 2007), door onvolwassenheid van de data warehouse processen (Sinha, 2012), door doorlooptijd van complexe data warehouse projecten (Takecian, 2014), maar ook in dure beheer- en onderhoudsprocessen. In zijn boek Business Intelligence: a managerial approach beschrijft E. Turban (2011) dat organisaties steeds complexer worden, maar ook de omgeving waarin ze opereren. Door globalisering, meer concurrentie, strengere regelgeving en hoge klanteisen, moeten organisaties voortdurend juiste, volledige en tijdige informatie opleveren. Zelfs Kimball (2016) merkte op dat: Het ontwerpen van een data warehouse is, anders dan een IT-taak, een combinatie van computertechnologie, cognitieve psychologie, business inhoud en politiek.

Alternatieve oplossingen: Met de opkomst van nieuwe technologieën en architecturen, zoals Big Data, Internet of Things (IoT), Data Lakes, Data Virtualisatie, Data Warehouse Automation, en flexibele dataopslag in de Cloud, neemt de interesse af en lijkt het huidige data warehouse grotendeels overbodig geworden (Serra, 2017). Dit blijkt ook uit figuur 1.2 die de interesse in data warehouse en big data vergelijkt.



Figuur 1.2 Zoekcriteria in Google de afgelopen 16 jaar. Bron: Google Trends

Data is tegenwoordig zo omvangrijk, gevarieerd en vluchtig (de 3 V's: Volume, Veriaty, Velocity), dat het geen zin heeft om alle data in een organisatie eerst te integreren en te modeleren. Een Big Data platform verzameld alle ruwe data eerst en pas als de data nodig is voor analyse of rapportage, wordt de data bewerkt en geïntegreerd. De nieuwe generatie analisten (data scientists en data-consumenten) bij de business, worden hiervoor steeds beter opgeleid.

Impact

Betekent dit het einde van het data warehouse zoals we dit kennen en de afgelopen 30 jaar geïmplementeerd en gebruikt hebben? Volgens vele opiniemakers wel, want “Niemand wil een data warehouse” van Rick van der Lans (2011) en “The Demise of the Data Warehouse” en “Is the traditional data warehouse dead?” van de vooraanstaand Microsoft blogger James Serra (2017), spreken wat dat betreft boekdelen.

Of is er toch nog hoop voor het data warehouse: “The Data Warehouse is dead – long live the Data Warehouse” (Raden, 2019) of een artikel van hét kennisinstituut uit Amerika TDWI – The Data Warehouse Institute (Schiff, 2019): “Long Live The Traditional Data Warehouse”, waarin

gesteld wordt dat de data warehouse-functies, zoals het integreren van kwalitatief hoogwaardige data op een tijdsafhankelijke manier, altijd nodig blijft.

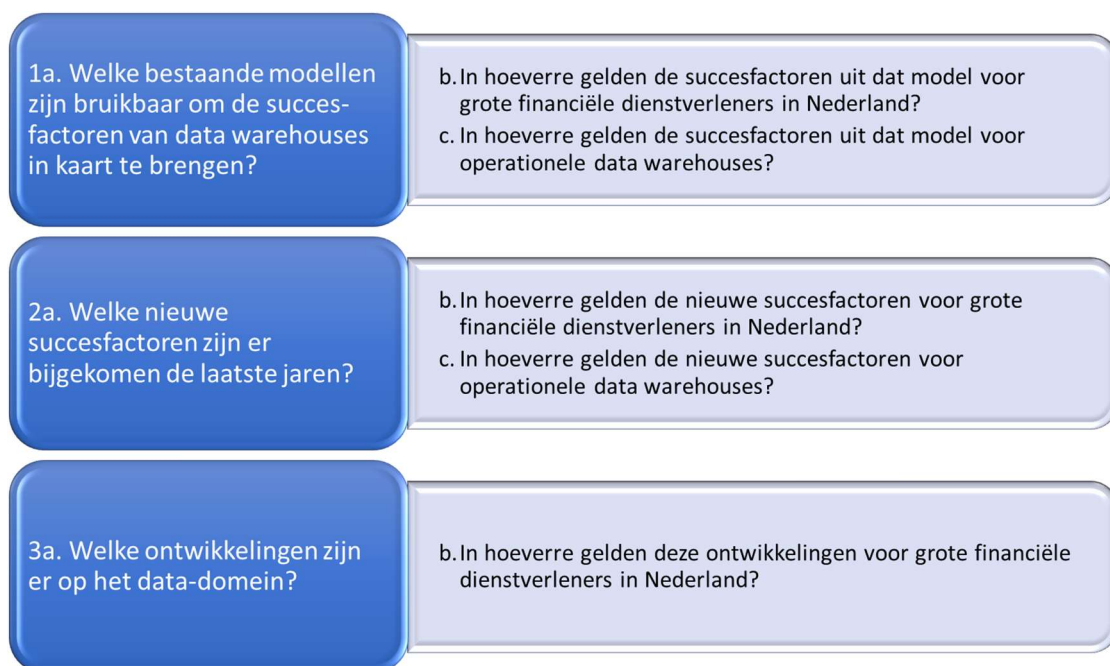
IBM heeft in 2015 onderzoek gedaan bij 90 organisaties wereldwijd (Difference, 2015) en stelde hen dezelfde vraag: Is het data warehouse dood? De conclusie was: Nee! 88% van de organisaties had één (23%) tot vijf data warehouses en alhoewel ontevreden (27%) over de omvang van de kosten en beheer in FTE, was 53% van de ondervraagde tevreden over hun data warehouse. Ze zagen de opkomst van bijvoorbeeld Big Data als belangrijk (64%) en een aanvulling (43%) op hun bestaande data warehouse.

1.3 Onderzoeksvraag

De hiervoor beschreven problemen hebben ook hun impact op huidige operationele data warehouse omgevingen in grote organisaties in Nederland, zoals banken en verzekeringsmaatschappijen. Maar wat is deze impact? In deze thesis wordt onderzocht of de in de literatuur beschreven succesfactoren van een data warehouse nog valide zijn of dat er nieuwe succesfactoren bijgekomen zijn de laatste jaren? De onderzoeksvraag luidt daarom:

Wat zijn de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland?

De deelvragen die gezamenlijk antwoord geven op de onderzoeksvraag zijn:



Figuur 1.3 Deelvragen data warehouse succesfactoren

Als we de onderzoeksvraag categoriseren naar de indeling van H. Oost (2016), dan valt dit onder: Verklaren (Explain): Het verklaren van een probleem door het opstellen van een theorie of model. De methode die hierbij vaak gebruikt wordt is de Theory Building van KM Eisenhardt (1989) of Theory Testing. Omdat er de nadruk ligt op bestaande data warehouse succesmodellen (zie literatuurstudie hoofdstuk 2), is dit onderzoek deductief van aard en wordt Theory Testing toegepast.

Het begrip kritische **succesfactoren** is beschreven door Rockart en Bullen (1981) en betekent: De factoren die beslissend zijn voor het al dan niet behalen van een vooraf gesteld doel. Om het doel te behalen ("succes") zijn bepaalde factoren een noodzakelijke of onderscheidende voorwaarde ("kritisch"). Het doel van een data warehouse is om de juiste data te bezorgen om besluitvorming op operationeel, tactisch en strategisch niveau te ondersteunen (Morris, 2013).

De definitie van een **data warehouse** is in de loop van de tijd geëvolueerd, zoals in paragraaf 1.1 staat beschreven, maar het primaire doel bleef gelijk: het ondersteunen van besluitvorming in een organisatie.

Alhoewel er meerdere omschrijvingen bestaan voor het begrip **grote financiële dienstverleners**, is gekozen voor de grootste groepen die het CBS hanteert: 1.000 tot 2.000 werknemers en meer dan 2.000 werknemers (CBS, 2020). De gehele populatie wordt beschreven in paragraaf 3.2.

De nadruk van de onderzoeksvraag ligt op **operationele** data warehouses. Operationeel wil zeggen dat deze systemen al vaak vele jaren actief draaien, soms wel 20 jaar, en een functie vervullen binnen de organisatie. Operationele data warehouses kennen andere uitdagingen en succesfactoren, dan nieuw te implementeren data warehouses. Banken en verzekeraars zijn van oudsher grote dataverwerkers (Cooper, 2000; Ramamurthy, 2007). Daarmee lopen ze vaak voorop als het gaat om IT en innovatie en dat betekent dat ze inmiddels veel ervaring hebben opgedaan met data warehouses.

1.4 Afbakening en scope

Business Intelligence

Het data warehouse maakt onderdeel uit van het gehele proces van Business Intelligence (BI) en ondersteunt de rapportage- en analysedoelstellingen van een organisatie. Vaak worden de begrippen data warehousing en BI door elkaar gebruikt. Een vrij recente definitie van BI is door Isik et al. (2013, pp. 13-14) beschreven:

“A system comprised of both technical and organizational elements that presents its users with historical information for analysis to enable effective decision making and management support, with the overall purpose of increasing organizational performance”.

Hierbij wordt met name het organisatorische aspect belicht als onderscheidend element ten opzichte van het technische: De dataverzameling in het data warehouse. Volgens Yeoh et al. (2008) zijn BI-systemen gelijk aan data warehouses, maar worden deze aangevuld met rapportage- en analysetoepassingen. Ook Chaudhuri et al. (2011) delen deze mening. Zij geven aan dat de data voor een BI-taak gehaald wordt uit een repository genaamd: Data warehouse. Veel onderzoek is gedaan naar BI-tools en de mate waarin ze ondersteuning bieden aan het proces van besluitvorming. Dit valt echter buiten scope van deze thesis.

Grote financiële organisaties

Grote organisaties zijn sterk in het mobiliseren van resources (geld en arbeid) om innovatie te bewerkstelligen (Ramamurthy, 2007). In Nederland waren het de grote banken en verzekeraars die begin deze eeuw de eerste innovatieve stappen zetten naar data warehouse succes. Vanwege het complexe datalandschap van grote organisaties is er meer nut en noodzaak voor een data warehouse (Winter, 2001). Een grote organisatie heeft ook meer business afdelingen en functionarissen die verantwoordelijk zijn, en eigenaarschap kunnen nemen voor de data (Winter, 2001). Dit is van belang voor data warehouse adoptie. Er is dus een positieve relatie tussen organisatiegrootte en adoptie en succes (Ramamurthy, 2007).

In Nederland is daarom gekeken naar de 25 grootste organisatie in de financiële wereld (CBS, 2020). In paragraaf 3.2 wordt hier verder op ingegaan.

1.5 Aanpak

Het antwoord op de 8 deelvragen moet een duidelijk antwoord geven op de onderzoeksvraag. Hieronder wordt opgesomd hoe de deelvragen worden beantwoord.

Literatuurstudie – Deelvragen 1a en 2a

Na het opkomende succes van data warehouse oplossingen in de jaren 90, is er in de jaren nul van deze eeuw veel onderzoek gedaan naar, en gepubliceerd over de succesfactoren.

1a. Er zijn drie kwantitatieve onderzoeken gevonden die ieder hun data warehouse succesmodel publiceerden. Het zijn de kwantitatieve modellen van Wixom et al. (2001), Hwang et al. (2008) en Ramamurty et al. (2007). De drie modellen zullen in dit onderzoek geïntegreerd worden tot één model. Dit samengestelde model is de theorie waarop deze thesis verder gaat.

2a. Door middel van literatuuronderzoek wordt gezocht naar de succesfactoren van de laatste 10 jaar. Dit moet de actualiteit van de succesfactoren verbeteren.

Enquête – Deelvragen 1b, 1c, 2b en 2c

Omdat de succesmodellen uit Amerika komen in een tijd dat veel nieuwe data warehouses geïmplementeerd werden, is een belangrijke vraag of deze succesfactoren ook gelden in Nederland, bij grote financiële organisaties, die al jaren één of meerdere data warehouses operationeel hebben. Het onderzoek-ontwerp (Research Design) is kwalitatieve Case Study Surveys (Chmiliar, 2012). Dit betekent dat het samengestelde model getoetst wordt door enquêtes met relevante personen en organisaties uit de doelgroep.

Interview – Deelvragen 3a en 3b

Door een aantal relevante personen te interviewen, zal inzicht verkregen worden welke ontwikkelingen er spelen op het data-domein in hun organisatie. Zijn dat aanvullingen of bedreigingen voor het operationele data warehouse? De informatie uit de interviews wordt geanalyseerd en gecombineerd met de antwoorden van een aantal open vragen uit de enquête.

In hoofdstuk 3 wordt de methodische aanpak verder toegelicht.

1.6 Contributie en toegevoegde waarde

Dit onderzoek heeft wetenschappelijke en praktische waarde, zoals hierna wordt toegelicht.

Praktische waarde

Zoals uit veel recente publicaties te lezen valt, staan we aan de vooravond van grote disruptieve veranderingen op het gebied van data en informatie. Maar hoever gaan grote Nederlandse organisaties hierin mee? Dit onderzoek kan inzicht geven in welke factoren het meeste invloed hebben op het succes van het data warehouse of het uitblijven daarvan. Nederlandse organisaties kunnen een zelf-assessment doen met behulp van de vragenlijst om te ontdekken welke factoren in hun organisatie onderbelicht zijn gebleven. Door gericht te investeren in deze factoren kan het hele datalandschap, inclusief het data warehouse, een positieve stimulans krijgen. Hier zullen alle eindgebruikers van informatiesystemen baat bij hebben en hun werk efficiënter en effectiever uit kunnen voeren, wat zal leiden tot beter onderbouwde besluitvorming.

Als bestaande succesmodellen vergeleken worden met de actuele versie uit dit onderzoek, is het dan voor organisaties interessant om te blijven investeren in hun huidige of toekomstige data-, rapportage- en analyse-systemen? Dit onderzoek kan aantonen dat succes steeds lastiger te realiseren is en dat er meer en betere alternatieven zijn voor een data warehouse omgeving. De uitkomst van dit onderzoek is met name interessant voor IT-managers en -leiders (CIO) en kan impact hebben op toekomstige data- en IT-plannen en -bestedingen binnen hun organisatie voor de komende jaren.

Academische waarde

Doel van dit onderzoek is de bestaande modellen van data warehouse succesfactoren te toetsen. Zijn er nieuwe factoren, die tegenwoordig het succes van een data warehouse bepalen? Kan er vanuit andere modellen één de-facto standaard gemaakt worden? In vervolgstudies kan het actuele model toegepast worden op andere data-systemen, zoals: Data Science, Big Data Analytics, Data Lakes, etc. Ook kan het model verder getoetst en verfijnd worden of kan er specifiek ingezoomd worden op één of meerdere factoren. De mogelijk nieuwe factoren die gevonden worden kunnen onderwerp zijn voor vervolgstudie.

Data warehousing als onderdeel van Business Intelligence, als onderdeel van ICT, is een serieus vakgebied wat meer aandacht verdiend vanuit de academische wereld.

2 Literatuurstudie

2.1 Inleiding

Al sinds de jaren 90 wordt er veel gepubliceerd over data warehouses en de nut en noodzaak daarvan. Het betreft vaak commerciële artikelen die gebruikt worden door IT-consultancy bureaus. Als wordt gekeken naar peer-reviewde wetenschappelijke artikelen die inductief of deductief te werk gaan, dan is de selectie kleiner. Bovendien wordt hier vaak verwezen naar dezelfde auteurs en naar elkaar, wat de academische wereld van data warehousing een stuk kleiner (en overzichtelijker) maakt.

De opbouw van dit hoofdstuk is 2-ledig.

In paragraaf 2.2 wordt gekeken naar een aantal **bestaande** data warehouse succes-modellen uit het begin van deze eeuw, toen data warehouses sterk in opkomst kwamen. Naast veel kwalitatieve publicaties zijn er drie kwantitatieve onderzoeken gevonden die elk een succesmodel beschrijven. Deze drie modellen worden kwalitatief geïntegreerd tot één samengesteld data warehouse succesmodel.

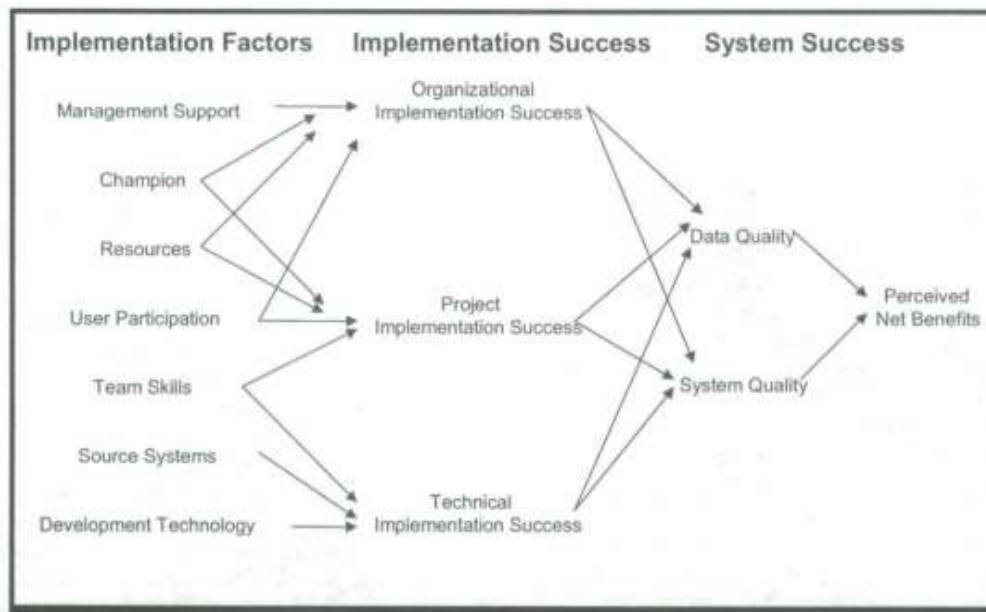
In paragraaf 2.3 wordt gekeken naar meer recente artikelen die het samengestelde model kunnen beïnvloeden of **veranderen**. De onderzoeksvraag gaat over operationele data warehouse succesfactoren, dus bij dit tweede deel wordt de nadruk gelegd op succesverhalen van (langdurig) bestaande en actieve data warehouse systemen. Zijn de hierbij gevonden factoren actueel en relevant en welke onderbouwde keuzes zijn er gemaakt? De onderzoeksvraag richt zich op Nederlandse organisatie, maar er is geen literatuur gevonden die zich hierop richt. Dit Nederlandse gedeelte van de onderzoeksvraag komt later aan bod tijdens de enquêtes en interviews.

2.2 Bestaande data warehouse modellen

Naast veel kwalitatieve publicaties zijn er drie kwantitatieve onderzoeken gevonden die een succesmodel beschrijven. De onderzoeken hebben plaatsgevonden in 2001, 2007 en 2008 en hun lijst met referenties is groot. Hun modellen zijn gebaseerd op een veelvoud aan interviews, onderzoeken en publicaties, waarbij ze meetwaarden hebben verzameld die de basis waren voor de succesfactoren in het model. In de volgende paragrafen worden de modellen uitgewerkt en toegelicht.

2.2.1 Empirisch model data warehouse succesfactoren

Barbara Wixom en Hugh Watson claimen in 2001 al (2001) dat ze als eersten wetenschappelijk en empirisch onderzoek gedaan hebben naar data warehouse implementatie succesfactoren. Gebaseerd op een literatuurstudie van voornamelijk IT-implementaties, een enquête onder 126 deelnemers van 111 bedrijven en 10 interviews met data warehouse experts, werd het volgende model gemaakt:



Figuur 2.1 Research model for data warehousing success van Wixom et al. (2001)

In hun onderzoek geven Wixom et al. (2001) aan dat er 3 dimensies zijn die data warehouse succes bepalen:

- Datakwaliteit: De data in het data warehouse moet van hoge kwaliteit zijn om besluitvorming te kunnen ondersteunen. De meest belangrijke aspecten hierbij zijn: juistheid, volledigheid en consistentie van data
- Systeemkwaliteit: Dit laat zich vertalen als de performance van een data warehouse, maar ook de betrouwbaarheid en beschikbaarheid van het systeem
- Door de business waargenomen voordelen gelden voor individuen, afdelingen, maar ook voor de gehele organisatie. De voordelen bestaan uit een beter begrip en context van de te nemen beslissing en de snelheid en frequentie waarmee beslissingen genomen kunnen worden.

Om bovenstaande succesfactoren positief te beïnvloeden moet de data warehouse implementatie op 3 vlakken succesvol zijn:

- Organisatie-implementatie: Zeker in 2001 was het belangrijk dat de business veranderde in hun manier van werken met de komst van automatisch beschikbaar gestelde data uit een data warehouse. De onderwerpen die hier genoemd worden zijn: nieuwe processen en functies, data-eigenaarschap en data-toegang & security
- Project-implementatie: Hier vermelden Wixom et al. (2001) weinig verschil met reguliere IT-projecten. Implementatiesucces wordt gemeten aan de hand van doelen, scope, tijd, budget en kwaliteit
- Technische implementatie: Hier onderscheidt een data warehouse implementatie zich wel van een reguliere IT-implementatie. Door grote volumes aan data te verwerken op (relatief) nieuwe technologieën heeft dit een extra impact op de infrastructuur en het aantal technische issues die opgelost moeten worden

De volgende factoren hebben een positieve relatie met implementatiesucces:

- Management support: Zonder sponsoring en draagvlak vanuit het management kan er geen organisatie-implementatie plaatsvinden
- Champion-rol die het project overziet, promoot, dagelijks betrokken is en alle stakeholders informeert
- Resources, zoals budget (euro's) en medewerkers (uren)
- Eindgebruikersparticipatie en -betrokkenheid waardoor de requirements helder zijn (en blijven) en de communicatie beter is
- Team skills hebben een positieve relatie tot het projectimplementatie- en technische implementatiesucces.

- Bronsystemen moeten betrokken zijn en door standaardisatie van hun aanlevering heeft dit een positief effect op het technische implementatiesucces
- Development tools als verzamelaar voor hardware, software en methoden, zijn noodzakelijk om effectief en efficiënt te werk te gaan

Wixom et al. (2001) concluderen dat data warehouse implementaties sterk lijken op generiek IT-implementaties, waarbij alleen de factoren datakwaliteit en systeemkwaliteit onderscheidend zijn.

De verzamelde meetwaarden die in het onderzoek beschreven staan zijn als volgt:

#	Factor	Meetwaarde Wixom et al.
1	Waargenomen voordelen	Verbeterde productiviteit
2	Waargenomen voordelen	Beter begrip van de context van de besluitvorming
3	Waargenomen voordelen	Snelle en makkelijke toegang tot data
4	Waargenomen voordelen	Verbeterde business processen
5	Datakwaliteit	De kwaliteit van de data beschikbaar in het data warehouse
6	Datakwaliteit	Juistheid van data
7	Datakwaliteit	Volledigheid van data
8	Datakwaliteit	Consistentie van data
9	Systeemkwaliteit	Flexibiliteit van het systeem
10	Systeemkwaliteit	Response tijd van het systeem
11	Systeemkwaliteit	Systeem is betrouwbaar (dit wordt niet verder toegelicht)
12	Systeemkwaliteit	Integratiecapaciteit van het systeem
13	Organisatie implementatie	Het systeem is opgeleverd en geaccepteerd door de business en geïntegreerd in de werkprocessen.
14	Project implementatie	Het projectteam is kundig en goed gemanaged
15	Technische implementatie	De technische complexiteit is groot door de hoeveelheid verschillende bronsystemen en grote datavolumes die initieel moeten worden ingeladen en beheerd
16	Management ondersteuning	Management ondersteuning en sponsoring is de belangrijkste succesfactor
17	Champion-rol	Een "champion" is aanwezig om het systeem te promoten, te ondersteunen en informatie te geven
18	Resources	Kundig data warehouse development team aanwezig
19	Resources	Er is voldoende geld, mensen en tijd (resources) aanwezig om het data warehouse te ontwerpen en te implementeren
20	Gebruikersbetrokkenheid	Betrokkenheid en deelname van eindgebruikers vanuit de business zorgt voor betere communicatie over requirements
21	Team kennis en kunde	Goede communicatie tussen developers en business eindgebruikers
22	Team kennis en kunde	Resources zijn nodig om mijlpalen binnen projecten op te leveren
23	Bronsystemen	Data uit bronsystemen moet hoogwaardig van kwaliteit zijn en gestandaardiseerd om een data warehouse implementatie succesvol te laten verlopen
24	Bronsystemen	Data Management realiseert significante voordelen voor de hele organisatie
25	Ontwikkelomgeving	Een goede technologische ontwikkelomgeving behelst hardware, software, methodes en heeft een positief effect op efficiency en effectiviteit

Figuur 2.2 Succesfactoren met hun meetwaarden. Bron: Wixom et al. (2001)

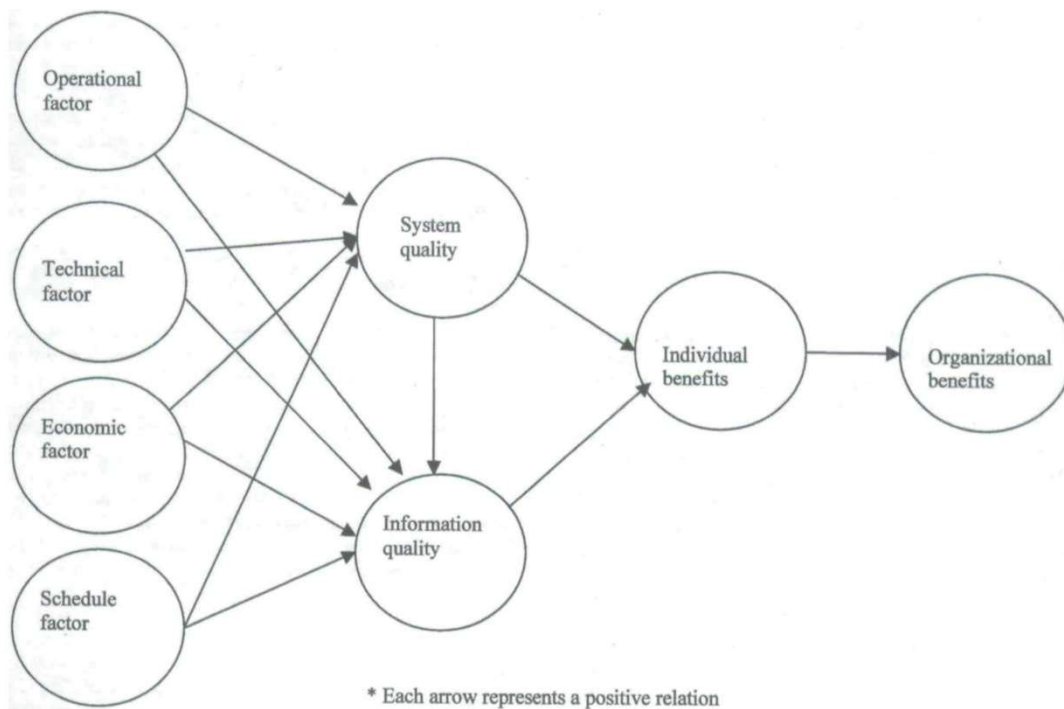
Gebruikerstevredenheid als meetwaarde is door Wixom et al. (2001) genoemd, maar bewust niet geselecteerd door hen, omdat een data warehouse geen applicatie is waar eindgebruikers op actief zijn.

2.2.2 Model data warehouse succesfactoren

Het meest recente peer-reviewde succesmodel is het onderzoek uit 2008, waarbij in een samenwerkingsverband tussen de Butler universiteit van Indianapolis en de Michigan University de onderzoekers Mark Hwang en Hongjiang Xu (2008) in hun proefschrift "A Structural Model of Data Warehousing Success" een research model beschreven hebben met factoren die effect hebben op data warehouse succes.

Naast een uitgebreide kwalitatieve literatuurstudie, waarin vele tientallen artikelen uit de periode 1990 – 2008 input hebben geleverd, is er een model tot stand gekomen vanuit kwantitatief onderzoek. Een groot aantal enquêtes (N=444) en artikelen over data warehousing hebben hiervoor gediend als input. Data is verzameld en geïnterpreteerd door middel van

statistische methodieken (oa Partial Least Square – PLS) en heeft geresulteerd in een regressiemodel met data warehouse succesfactoren en hun relaties.



Figuur 2.3 Research model for data warehousing success. Bron: Hwang et al. (2008)

De directe factoren in dit model die de basis vormen voor een succesvol data warehouse, zijn:

1. Organisatorische voordelen:

De organisatorische voordelen zijn feitelijk de doelstelling die nagestreefd en behaald worden met het data warehouse. Op strategisch niveau betekent dit ondersteuning van besluitvorming om de organisatorische doelen, zoals verhogen omzet / winst, verlagen van kosten, betere concurrentiepositie. Er is ook voordeel te behalen aan het verbeteren of ondersteunen van bedrijfsprocessen;

2. Individuele voordelen:

De individuele voordelen die met een data warehouse zijn te behalen zitten meer op operationeel niveau. Bepaalde detail-rapporten of lijstwerk ondersteunen de dagelijkse werkzaamheden en dragen zorg voor een betere individuele productiviteit;

3. Systeemkwaliteit:

Bij dit onderwerp gaat het over de kwaliteit van de hardware (infrastructuur), de database en de query tools rond het data warehouse. Zijn deze zaken goed ingericht dan heeft dat een directe positieve bijdrage aan het gebruik, het gebruiksgemak en daarmee het succes van een data warehouse omgeving. Ook documentatie en security is een belangrijke factor hierbij;

4. Informatie- of Datakwaliteit:

Een onderscheid tussen een data warehouse en een algemeen IT-systeem is de datakwaliteit. Deze moet foutvrij, actueel en informatief zijn, anders wordt het data warehouse niet gebruikt. Of erger nog, de beslissingen op strategische en tactisch niveau worden gemaakt op basis van onjuiste informatie, met alle gevolgen van dien.

Bovenstaande factoren worden beïnvloed door de volgende indirecte factoren:

1. Operationele factoren:

Hier gaat het om de business doelstellingen en betrokkenheid. Een data warehouse moet voortdurend in lijn zijn met de snel veranderende business processen en strategische doelstellingen. Is de directie of het (top)management voldoende betrokken en leveren ze tijdig resources in de vorm van mensen en middelen?

2. Technische factoren:

De hardware en software moet actueel zijn en gesupport worden door de leveranciers. De IT-medewerkers hebben voldoende technische kennis en kunde. Specifiek voor een data warehouse omgeving is de kwaliteit van de data van de aanleverende systemen. Deze is vaak onvoldoende en hierop moet adequaat op gereageerd worden. Zijn er tools in gebruik rond het proces van Data Management, zoals metadata, masterdata en data profiling?

3. Planningsfactoren:

Zoals elk IT-systeem is planning en doorlooptijd ook bij een data warehouse van belang. Een data warehouse omgeving moet snel kunnen inspelen op continue wijzigende vragen vanuit de business. Gebeurt dit door een agile werkwijze of waterval projecten? Zijn er externe ontwikkelingen die de planning van de data warehouse omgeving beïnvloeden?

4. Economische (en organisatorische) factoren:

Is er voldoende funding voor beheer en verbetering? Worden de benefits van een data warehouse voortdurend kwalitatief en kwantitatief gemeten? Wordt het binnen de organisatie gezien als een innovatief instrument of als (IT) kostenpost?

Relaties

Alle indirecte factoren in het model hebben een positieve relatie tot de directe factoren: Systeem- en informatiekwaliteit. Deze twee directe factoren op hun beurt hebben een positieve relatie met de factor: individuele voordelen die een positieve relatie heeft met de factor: organisatorisch voordelen. Volgens het regressiemodel van Hwang en Xu is de relatie tussen deze laatste twee het sterkst.

Opvallend is de eenzijdige relatie tussen systeem- en informatiekwaliteit. In het artikel wordt aangegeven dat, een systeem met de gewenste karakteristieken positief bijdraagt aan het eindproduct informatie, wat bijdraagt aan de individuele voordelen. Andersom wordt niet vermeld, want goede informatie draagt blijkbaar niet bij aan een beter systeem. Dit is discutabel als gekeken wordt vanuit het standpunt van een business gebruiker, die een data warehouse beoordeeld als één geheel: goede informatie uit een betrouwbaar systeem.

De verzamelde meetwaarden die in het onderzoek beschreven staan zijn als volgt:

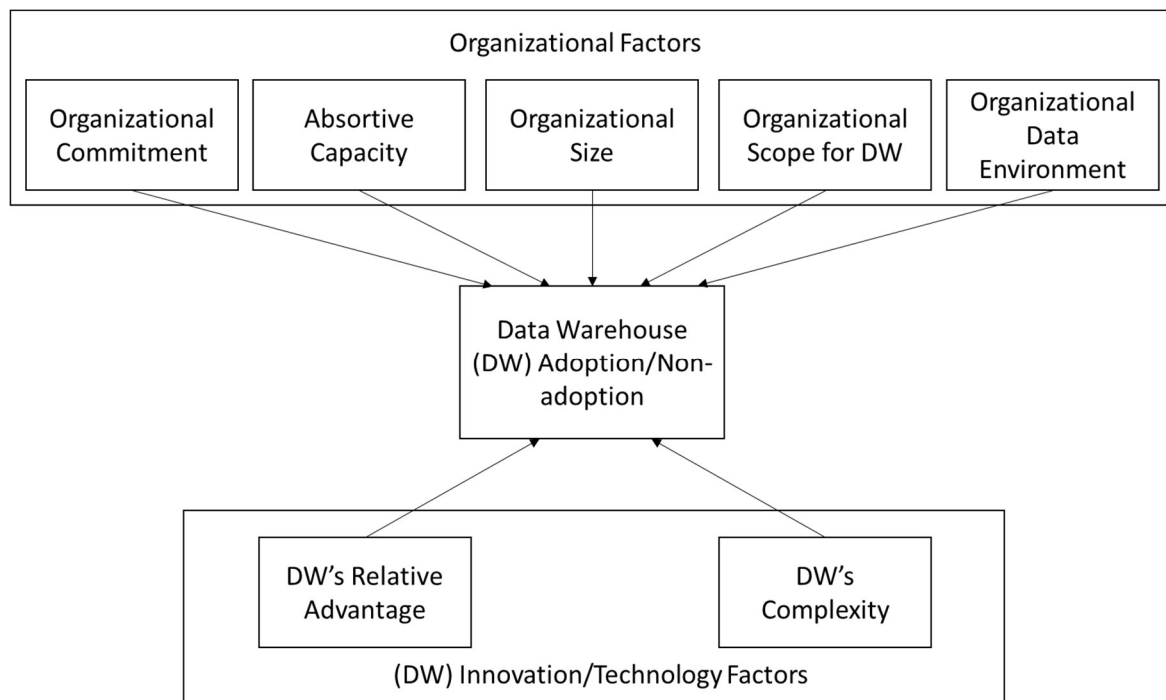
#	Factor	Meetwaarde Hwang et al.
1	Organisatorische voordelen	Verbeterde business processen
2	Organisatorische voordelen	Verbeterde concurrentiepositie, door verlagen van de kosten en/of verhogen van de winst
3	Individuele voordelen	Verbeterde productiviteit
4	Individuele voordelen	Betere besluitvorming
5	Systeemkwaliteit	Systeem is makkelijk in gebruik
6	Systeemkwaliteit	Snel en efficiënt ophalen van benodigde informatie
7	Systeemkwaliteit	Het systeem is schaalbaar, gestandaardiseerd en de security is geregeld
8	Informatiekwaliteit	Goede kwaliteit van informatie
9	Informatiekwaliteit	Meer informatie beschikbaar
10	Operationele factor	Duidelijk gedefinieerde business behoeften en voordelen
11	Operationele factor	Ondersteuning op directieniveau
12	Operationele factor	Betrokkenheid en deelname van eindgebruikers vanuit de business
13	Technische factor	Datakwaliteit van de bronsystemen
14	Technische factor	Professionele ontwikkelomgeving
15	Technische factor	Adequate IT management en (ingehuurde) consultants
16	Technische factor	Project Management en werken in teamverband
17	Economische factor	Adequate financiering / budgettering
18	Economische factor	Meetbare business voordelen (kosten-baten analyse), vooral bij uitbreidingen op het data warehouse
19	Planningsfactor	Praktische planning voor implementaties / releases
20	Planningsfactor	Juiste planning en scoping van de projecten

Figuur 2.4 Succesfactoren met hun meetwaarden. Bron: Hwang et al. (2008)

2.2.3 Model data warehouse adoptie

Succes kan gemeten worden in de mate waarin een data warehouse gebruikt wordt. De adoptie van een data warehouse is uitgebreid onderzocht door de heren Ramamurthy, Sen en Sinha in hun publicatie in 2007 (2007). In hun peer-reviewde onderzoek vanuit de universiteiten van Wisconsin-Milwaukee en Texas (Verenigde Staten) hebben de onderzoekers een data warehouse adoptiemodel beschreven.

De basis van hun onderzoek is een enquête die verspreid is over 2.500 organisaties in de Verenigde Staten. In totaal hebben 196 organisaties gerespondeerd en die resultaten zijn verzameld en geïnterpreteerd door middel van statistische methodieken (Linear Structural Relations – LISREL) wat resulteerde in een regressiemodel met data warehouse adoptiefactoren en hun relaties.



Figuur 2.5 Research model for data warehouse adoption. Bron: Ramamurthy et al. (2007)

Het onderzoek focust zich op data warehouse adoptie en welke factoren hierbij spelen. Er wordt een bestaand model vanuit de IT toegepast op Decision Support Systems (DSS), waar een data warehouse ook onder valt. Er zijn 7 factoren gevonden (5 organisatorische en 2 technische). Een data warehouse wordt gezien als een innovatieve en infrastructurele oplossing, maar daarmee wordt de software en tooling niet meegenomen. De 7 factoren die bepalend zijn voor data warehouse adoptie, zijn:

1. Organisatie commitment

Onder commitment vallen twee aspecten: Ondersteuning vanuit de directie (senior management) en betrokkenheid van de belangrijkste stakeholders. Beiden hebben een noodzakelijke positieve relatie met data warehouse adoptie

2. Absorptievermogen

Hieronder wordt verstaan, het vermogen van een organisatie om zijn beschikbare kennis en kunde in te zetten. Maar ook om in te zien wat de waarde en noodzaak is van een data warehouse als innovatieve oplossing

3. Organisatiegrootte

Kleine organisaties zijn efficiënter in het implementeren en accepteren van innovatie (hier valt een data warehouse onder) dan grote organisaties. Grote organisaties zijn daarentegen sterk in het mobiliseren van resources (geld en arbeid) om innovatie te bewerkstelligen. Volgens de onderzoekers is dit laatste van belang voor data warehouse adoptie dus er is een positieve relatie tussen organisatiegrootte en adoptie en succes

4. Organisatie scope van het data warehouse

De scope en voordelen van een IT-innovatie moeten goed gedefinieerd zijn om data warehouse adoptiesucces te vergroten

5. Organisatie data omgeving

Data Management met al zijn facetten worden onderkent als voorwaarde voor data warehouse adoptie. Met name de datakwaliteit van de bestaande IT-systemen worden hier genoemd

6. Data warehouse voordelen

Uitgangspunt hierbij is dat innovatie (data warehouse) altijd een waargenomen verbetering is van hetgeen vervangen wordt. Indien de voordelen van een data warehouse voor de organisatie beschreven en bekend zijn, heeft dat een positieve relatie tot data warehouse adoptie

7. Data warehouse complexiteit

Een hoge mate van complexiteit van een data warehouse heeft een negatieve relatie tot het begrijpen en gebruiken ervan en daarmee ook tot data warehouse adoptie

De uitkomst van dit onderzoek heeft aangetoond dat de factoren 4 en 5 minder van invloed zijn, dan de andere vijf adoptiefactoren. Het onderzoek geeft aan dat vervolgstudie noodzakelijk is op de onderwerpen: data warehouse flexibiliteit, response-snelheid, standaardisatie, consistentie, integratie, beheer en datakwaliteit.

De verzamelde meetwaarden die in het onderzoek beschreven worden zijn als volgt:

#	Factor	Meetwaarde Ramamurthy et al.
1	Organisatie commitment	Senior management ondersteuning, betrokkenheid en buy-in
2	Organisatie commitment	Eindgebruikers en stakeholders moeten betrokken zijn en deelnemen aan innovatie in een organisatie
3	Organisatie commitment	Senior management heeft een data strategische visie ontwikkeld voor het data-domein
4	Organisatie commitment	De data en datakwaliteit in een organisatie moet goed worden beheerd en gemanaged
5	Absorptievermogen	Absorptievermogen: Zijn de betrokkenen van een data warehouse in staat en beschikbaar om hun kennis en kunde in te zetten?
6	Absorptievermogen	Data warehouse kan zorgen voor een verandering van bestaande business processen.
7	Absorptievermogen	Absorptie van informatie is noodzakelijk ter ondersteuning van operationele en strategische activiteiten en verbeterde organisatorische prestaties
8	Absorptievermogen	Absorptievermogen is voorwaardelijk voor een flexibele organisatie die adequaat reageert op een veranderde markt
9	Absorptievermogen	Is de organisatie innovatief genoeg en wordt daarbij de toegevoegde waarde van een data warehouse onderkend?
10	Organisatiegrootte	Een grote organisatie heeft meer innovatieve slagkracht in de vorm van mensen (kennis en kunde) en middelen (geld)
11	Organisatiegrootte	Een grote organisatie heeft meer baat bij een data warehouse, omdat deze efficiënter en effectiever gebruik kunnen maken, tegen een relatief lagere investering
12	Scope	Hoe groter het IT-landschap (infra en applicaties), hoe meer kans op data warehouse adoptie
13	Organisatie data omgeving	Een goed gemanaged data-omgeving bevordert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
14	Organisatie data omgeving	Data (resource) management biedt veel voordelen voor een organisatie
15	Organisatie data omgeving	Data (resource) management zorgt voor minder fouten en verbetert de toegankelijkheid van data
16	Data warehouse voordelen	Ondersteunen van effectieve besluitvorming door rapportage
17	Data warehouse voordelen	Faciliteren van OLAP / data-analyse
18	Data warehouse voordelen	Een data warehouse faciliteert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
19	Complexiteit	Complexiteit van een data warehouse (in de ogen van de gebruiker) is een bedreiging voor de adoptie ervan

Figuur 2.6 Succesfactoren met hun meetwaarden. Bron: Ramamurthy et al. (2007)

2.2.4 Samengesteld model succesfactoren

In deze paragraaf worden de drie bestaande data warehouse modellen samengevoegd tot één geïntegreerd data warehouse succesmodel.

De samenvoeging komt tot stand door de meetwaarden (of variabelen) te analyseren van elke factor die vermeld wordt in de drie modellen. Die meetwaarden worden vergeleken en indien van toepassing, samengevoegd. Dit zal leiden tot één samengesteld model, waarop de vragenlijst gebaseerd kan worden voor het vervolg van deze thesis.

Leidend model

Wixom et al. (2001) heeft de basis gelegd in 2001 door de data warehouse succesfactoren te vergelijken met reguliere IT-implementaties en vond grote overeenkomsten en een paar aanvullingen: datakwaliteit en systeemkwaliteit. Ook Hwang et al. (2008) in 2008 kwam tot dezelfde conclusie en gaf beide factoren een prominente plek in het succesmodel. Dit is deels te verklaren doordat de resultaten van Wixom et al. meegenomen zijn in het onderzoek van Hwang et al. Omdat Hwang et al. meer focust op de operationele factoren van een data warehouse, is dit model beter toepasbaar bij de onderzoeksvraag in deze thesis.

Het adoptiemodel van Ramamurtha et al. (2007) in 2007 richt zich op de gebruikers van het datawarehouse en de adoptie ervan. Uitgangspunt hierbij is dat de adoptie en het gebruik van een data warehouse door de business eindgebruiker gelijk staat aan het succes ervan. Dit komt overeen met de doelen van een data warehouse: het ondersteunen van besluitvorming.

Samenvoegen

In drie stappen worden de modellen geïntegreerd tot één succesmodel:

Stap 1. Alle factoren en meetwaarden van de 3 modellen zijn verzameld en in Excel naast elkaar geplaatst, te beginnen bij het leidend model van Hwang et al. (2008);

Stap 2. De meetwaarden worden één voor één vergeleken, waarbij geldt: Overeenkomstige meetwaarden worden geïntegreerd. De overgebleven unieke meetwaarden krijgen ook een plaats in het samengestelde model bij een bestaande succesfactor. Op die manier gaat geen enkele meetwaarde verloren.

Dit leidt tot een totaalbeeld van factoren en meetwaardes in figuur 2.7.

#	Factor	Meetwaarde Hwang et al.	#	Factor	Meetwaarde Wixom et al.	#	Factor	Meetwaarde Ramamurthy et al.
H01	Organisatorische voordelen	Verbeterde business processen	W04	Waargenomen voordelen	Verbeterde business processen	R06	Absorptievermogen	Data warehouse kan zorgen voor een verandering van bestaande business processen.
H02	Organisatorische voordelen	Verbeterde concurrentiepositie, door verlagen van de kosten en/of verhogen van de winst				R08	Absorptievermogen	Absorptievermogen is voorwaardelijk voor een flexibele organisatie die adequaat reageert op een veranderde markt
H03	Individuele voordelen	Verbeterde productiviteit	W01	Waargenomen voordelen	Verbeterde productiviteit	R07	Absorptievermogen	Absorptie van informatie is noodzakelijk ter ondersteuning van operationele en strategische activiteiten en verbeterde organisatorische prestaties
H04	Individuele voordelen	Betere besluitvorming	W02	Waargenomen voordelen	Beter begrip van de context van de besluitvorming	R16	Data warehouse voordelen	Ondersteunen van effectieve besluitvorming door rapportage
						R17	Data warehouse voordelen	Faciliteren van OLAP / data-analyse
H05	Systeemkwaliteit	Systeem is makkelijk in gebruik	W09	Systeemkwaliteit	Flexibiliteit van het systeem	R19	Complexiteit (geen succes-, maar een faalfactor)	Complexiteit van een data warehouse (in de ogen van de gebruiker) is een bedreiging voor de adoptie ervan
H06	Systeemkwaliteit	Snel en efficiënt ophalen van benodigde informatie	W03	Waargenomen voordelen	Snelle en makkelijke toegang tot data	R19	Complexiteit (geen succes-, maar een faalfactor)	Complexiteit van een data warehouse (in de ogen van de gebruiker) is een bedreiging voor de adoptie ervan
H07	Systeemkwaliteit	Het systeem is schaalbaar, gestandaardiseerd en de security is geregeld	W10	Systeemkwaliteit	Response tijd van het systeem			
			W12	Systeemkwaliteit	Integratiecapaciteit van het systeem	R18	Data warehouse voordelen	Een data warehouse faciliteert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
			W11	Systeemkwaliteit	Systeem is betrouwbaar (dit wordt niet verder toegelicht)			
H08	Informatiekwaliteit	Goede kwaliteit van informatie	W05	Datakwaliteit	De kwaliteit van de data beschikbaar in het data warehouse	R15	Organisatie data omgeving	Data (resource) management zorgt voor minder fouten en verbetert de toegankelijkheid van data
H09	Informatiekwaliteit	Meer informatie beschikbaar	W07	Datakwaliteit	Volledigheid van data			
			W06	Datakwaliteit	Juistheid van data	R18	Data warehouse voordelen	Een data warehouse faciliteert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
			W08	Datakwaliteit	Consistentie van data	R18	Data warehouse voordelen	Een data warehouse faciliteert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
H10	Operationele factor	Duidelijk gedefinieerde business behoeften en voordelen	W20	Gebruikersbetrokkenheid	Betrokkenheid en deelname van eindgebruikers vanuit de business zorgt voor betere communicatie over requirements			
H11	Operationele factor	Ondersteuning op directieniveau	W16	Management ondersteuning	Management ondersteuning en sponsoring is de belangrijkste succesfactor	R01	Organisatie commitment	Senior management ondersteuning, betrokkenheid en buy-in
H12	Operationele factor	Betrokkenheid en deelname van eindgebruikers vanuit de business	W17	Champion-rol	Een "champion" is aanwezig om het systeem te promoten, te ondersteunen en informatie te geven	R02	Organisatie commitment	Eindgebruikers en stakeholders moeten betrokken zijn en deelnemen aan innovatie in een organisatie
			W21	Team kennis en kunde	Goede communicatie tussen developers en business eindgebruikers			
			W22	Team kennis en kunde	Resources zijn nodig om mijlpalen binnen projecten op te leveren			
			W13	Organisatie implementatie	Het systeem is opgeleverd en geaccepteerd door de business en geïntegreerd in de werkprocessen.			
						R03	Organisatie commitment	Senior management heeft een data strategische visie ontwikkeld voor het data-domein
H13	Technische factor	Datakwaliteit van de bronsystemen	W23	Bronsystemen	Data uit bronsystemen moet hoogwaardig van kwaliteit zijn en gestandaardiseerd om een data warehouse implementatie succesvol te laten verlopen	R13	Organisatie data omgeving	Een goed gemanaged data-omgeving bevordert data-integriteit, accuraatheid, security, standaardisatie en data-uitwisseling
			W24	Bronsystemen	Data Management realiseert significante voordelen voor de hele organisatie	R14	Organisatie data omgeving	Data (resource) management biedt veel voordelen voor een organisatie
						R04	Organisatie commitment	De data en datakwaliteit in een organisatie moet goed worden beheerd en gemanaged
H14	Technische factor	Professionele ontwikkelomgeving	W25	Ontwikkelomgeving	Een goede technologische ontwikkelomgeving behelst hardware, software, methodes en heeft een postief effect op efficiency en effectiviteit			
H15	Technische factor	Adequate IT management en (ingehuurde) consultants	W18	Resources	Kundig data warehouse development team aanwezig	R05	Absorptievermogen	Absorptievermogen: Zijn de betrokkenen van een data warehouse in staat en beschikbaar om hun kennis en kunde in te zetten?
H16	Technische factor	Project Management en werken in teamverband	W14	Project implementatie	Het projectteam is kundig en goed gemanaged			
			W15	Technische implementatie	De technische complexiteit is groot door de hoeveelheid verschillende bronsystemen en grote datavolumes die initieel moeten worden ingeladen en beheerd	R12	Scope	Hoe groter het IT-landschap (infra en applicaties), hoe meer kans op data warehouse adoptie
H17	Economische factor	Adequate financiering / budgettering	W19	Resources	Er is voldoende geld, mensen en tijd (resources) aanwezig om het data warehouse te ontwerpen en te implementeren	R10	Organisatiegrootte	Een grote organisatie heeft meer innovatieve slagkracht in de vorm van mensen (kennis en kunde) en middelen (geld)
H18	Economische factor	Meetbare business voordelen (kosten-baten analyse), vooral bij uitbreidingen op het data warehouse				R09	Absorptievermogen	Is de organisatie innovatief genoeg en wordt daarbij de toegevoegde waarde van een data warehouse onderkend?
						R11	Organisatiegrootte	Een grote organisatie heeft meer baat bij een data warehouse, omdat deze efficiënter en effectiever gebruik kunnen maken, tegen een relatief lagere investering
H19	Planningsfactor	Praktische planning voor implementaties / releases						
H20	Planningsfactor	Juiste planning en scoping van de projecten	W14	Project implementatie	Het projectteam is kundig en goed gemanaged			

Figuur 2.7 Samengevoegde data warehouse succesmodellen, gegroepeerd per succesfactor (kleur)

In het boek: Completing Your Qualitative Dissertation: A Roadmap From Beginning to End (Bloomberg, 2008), beschrijven Bloomberg & Volpe een methode om op een gestructureerde manier kwalitatieve data-analyse te doen. De verzamelde data worden als volgt georganiseerd:

- Stap 1. Categorieën (“Categories”) ontwikkelen;
- Stap 2. Beschrijvingen (“Descriptors”) maken per categorie.

Bovenstaande wordt in een “Data Summary Table” (dataverzameltabel) geplaatst, die gebruikt wordt voor kwalitatieve data-analyse. Dit is goed toe te passen op dit onderzoek, want de categorieën zijn de gevonden succesfactoren en de beschrijvingen de bijbehorende meetwaarden. De verzameltabel dient als input voor de vervolgstappen in dit onderzoek: literatuurstudie naar nieuwe factoren en meetwaarden, en het analyseren van enquête- en interview-resultaten.

Factor / Categorie	Nr	Meetwaarde / Beschrijving
Waargenomen voordelen	V1	Ondersteunen primaire, secundaire en innovatieve processen
	V2	Ondersteunen van strategische doelstelling
	V3	Verbeterde individuele productiviteit
	V4	Verbeterde besluitvorming op strategisch en tactisch niveau
	V5	Verbeterde besluitvorming op operationeel niveau
	V6	Ondersteunen van analyse-doeleinden
Systeemkwaliteit	S1	Makkelijk in het gebruik en gebruiksvriendelijk
	S2	Eenvoudig en efficiënt
	S3	Goede responsetijd door schaalbaarheid
	S4	Gestandaardiseerd, gedocumenteerd en onderhoudbaar
	S5	Security en autorisatie is goed geregeld
	S6	Integratie van meerdere bronsystemen
	S7	Goede beschikbaarheid
Datakwaliteit	D1	Hoogwaardige datakwaliteit
	D2	Voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens
	D3	Toereikende actualiteit van data (laadfrequentie en verversingsgraad)
	D4	Toegang tot geïntegreerde en consistente data
Organisatorische factoren	O1	Doelstellingen zijn bekend en voldoen aan de business behoeften
	O2	Management vanuit de business en IT is betrokken en ondersteunt
	O3	Business (eind)gebruikers zijn actief betrokken
	O4	Goede communicatie tussen IT en business
	O5	Resources zijn beschikbaar gesteld
	O6	Er is een visie en roadmap
Technische factoren	T1	Goede datakwaliteit aanleverende bronsystemen
	T2	Organisatorische data wordt voldoende beheerd
	T3	Goede ontwikkelomgeving
	T4	Kundig IT-personeel
	T5	Continue verbeteringen vanuit projecten
	T6	Veel data en bronsystemen, waardoor technisch complex
Economische factor	E1	Voldoende funding
	E2	Voordelen zijn bekend en worden financieel gemeten
	E3	Grootte organisatie
Planningsfactor	P1	Verbeterpunten zijn bekend en geregistreerd
	P2	Wijzigingen worden goed uitgevoerd en gemanaged

Figuur 2.8 Dataverzameltabel samengesteld uit bestaande data warehouse succesmodellen

Zoals aangegeven zijn door de samenvoeging geen meetwaarden verloren gegaan. Voorbeelden hiervan zijn:

- Ramamurthy et al. (2007) zijn met hun adoptiemodel meer gefocust op commitment vanuit de organisatie en noemt als enige de noodzaak van het hebben van een data-visie (O6);
- Hwang et al. (2008) hebben als enige in 2008 al meer interesse in wijzigingen en verbeterpunten van een data warehouse (P1 en P2);

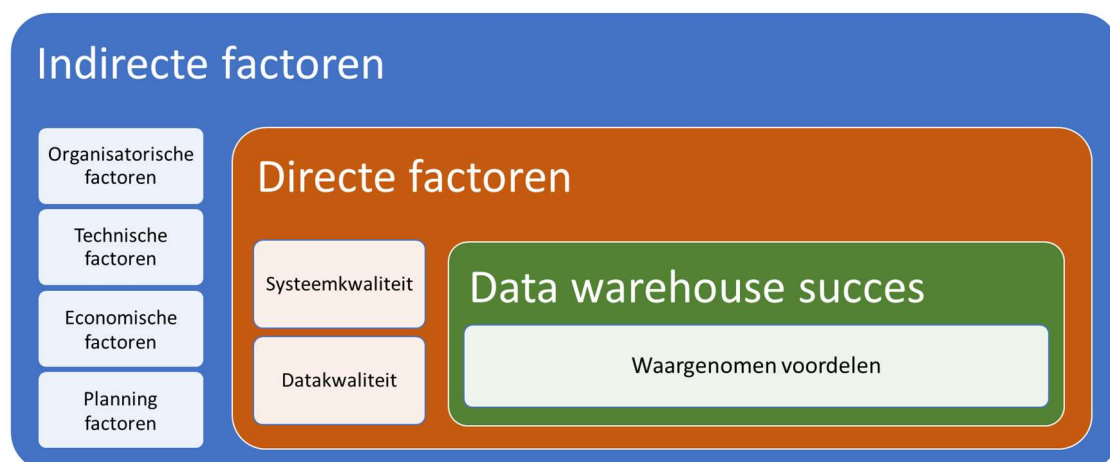
– Wixom et al. (2001) benoemen een kundig data warehouse development team (T4). Net als bij de evolutie van de data warehouse definities in paragraaf 1.1, zijn er in de loop van de jaren ook meetwaarden en succesfactoren bijgekomen en bijgesteld. Al deze factoren worden in het model meegenomen en getoetst bij Nederlandse financiële dienstverleners.

Stap 3. Wat valt op na een vergelijkingsanalyse van drie modellen:

- Individuele voordelen en organisatorische voordelen bij Hwang et al. (2008) hebben sterke overlap. Wixom et al. (2001) en Ramamurthy et al. (2007) maken dit onderscheid niet. In het verleden werd het data warehouse alleen gebruikt door een aantal data-savvy business eindgebruikers, waardoor dit onderscheid interessant was te melden. Tegenwoordig wordt gekeken naar enterprise data en de business processen en is de individu minder relevant. Ramamurthy et al. zegt hierover dat een data warehouse, met zijn beleid van data-integratie en architectuur-integratie, een sterke focus heeft op ondersteuning van besluitvorming en technologische innovatie binnen de gehele organisatie. Het is dus geen tool voor individuen of alleen de IT-organisatie.
 - o **Conclusie: de factoren individuele voordelen en organisatorische voordelen worden samengevoegd tot: Waargenomen voordelen**
- De meetwaarden behorend bij de Operationele factoren bij Hwang et al. (2008); komen sterk overeen met de meetwaarden bij de factor organisatie commitment van Ramamurthy et al. (2007). Ook de vergelijkbare meetwaarden bij Wixom et al. (2001) inzake betrokkenheid van de business, management ondersteuning, champion-rol, communicatie IT en business en acceptatie, hebben een sterk organisatorisch karakter.
 - o **Conclusie: de operationele factor wordt hernoemd naar: organisatorische factor**

De relaties tussen de factoren in het leidende model van Hwang et al. (2008) zijn allemaal positief gecorreleerd, dat wil zeggen dat elke indirecte factor een positieve invloed heeft op alle directe factoren, die vervolgens allebei positief bijdragen aan de waargenomen voordelen en daarmee aan het succes. Om die reden voegen de relaties geen toegevoegde waarde in het samengestelde model en worden ze daarom niet expliciet gemodelleerd.

Het nieuwe samengestelde model ziet er dan als volgt uit:



Figuur 2.9 Samengesteld model succesfactoren data warehouses

2.3 Recente operationele data warehouse succesfactoren

In deze paragraaf wordt in de literatuur gezocht naar nieuwe succesfactoren of meetwaarden van de laatste 10 jaar. Uitgangspunt van dit onderzoek is de dataverzameltabel (figuur 2.7) en het samengestelde succesmodel (figuur 2.8) op basis van de modellen van Wixom et al. (2001), Hwang et al. (2008) en Ramamurty et al. (2007). Belangrijkste vraag in dit deel van de literatuurstudie: Zijn er na 2008 factoren bijgekomen of afgevallen ten opzichte van het samengestelde model? Zijn er factoren die met name relevant zijn als we kijken naar (langdurig) operationele data warehouse successen?

Het doel van deze literatuurstudie is daarmee tweeledig:

1. De relevante publicaties bevestigen het samengestelde model;
2. De relevante publicaties beschrijven nieuwe meetwaarden die het succes van een operationeel data warehouse onderbouwen.

In bijlage 8.1 staan de criteria en zoektermen genoemd.

De onderzoeksvraag richt zich op Nederlandse organisatie, maar er is geen literatuur gevonden die zich hierop richt. Het Nederlandse gedeelte van de onderzoeksvraag komt later aan bod tijdens de enquêtes en interviews (hoofdstuk 3).

2.3.1 Bevestiging bestaande factoren

Over de doelen en functies van een data warehouse is de vakliteratuur het erg eens. Dit is sinds de introductie in de jaren 90 niet of nauwelijks veranderd. Veel artikelen gaan over de succesfactoren van een data warehouse-implementatie en niet over de succesfactoren van een (langlopend) operationeel data warehouse.

De resultaten van het literatuuronderzoek zijn uitgewerkt in bijlage 8.2. Zodra een artikel gevonden wordt en beschikbaar gesteld is, wordt deze beoordeeld op: Relevantie (bijlage 8.3) en worden de gevonden meetwaarden, functies en/of factoren uitgewerkt.

Van de 31 gevonden artikelen, zijn er:

- 5 niet gevonden, omdat de aanvraag nog loopt of omdat ervoor betaald moest worden;
 - 3 niet relevant geacht, omdat er geen meetwaarden / succesfactoren beschreven werden.
- De artikelen 12, 14 en 24 beschrijven de data warehouse succesmodellen die input zijn voor het samengestelde succesmodel in de vorige paragraaf. Deze worden hier niet (nogmaals) meegenomen.

In onderstaande tabel wordt gekeken naar de succesfactoren uit het samengestelde model van paragraaf 2.2.4, afgezet tegen de in de recente literatuur gevonden meetwaarden, die horen bij de factoren. Dit geeft een beeld hoe vaak de factoren genoemd worden en dat bevestigt de geldigheid en actualiteit ervan.

#	Factor	Artikelnummer(s)	Aantal
1	Waargenomen voordelen	Art.1, Art.3, Art.4, Art.5, Art.6, Art.7, Art.13, Art.15, Art.16, Art.20, Art.23, Art.26, Art.28, Art.30	14
2	Systeemkwaliteit	Art.1, Art.3, Art.5, Art.7, Art.13, Art.16, Art.22, Art.23, Art.26	9
3	Datakwaliteit	Art.5, Art.15, Art.16, Art.20	4
4	Organisatorische factor	Art.1, Art.4, Art.5, Art.6, Art.7, Art.16, Art.20, Art.21, Art.28	9
5	Technische factor	Art.1, Art.3, Art.4, Art.5, Art.6, Art.7, Art.16, Art.20, Art.21, Art.27, Art.28	11
6	Economische factor	Art.3, Art.4, Art.5, Art.6	4
7	Planningsfactoren	Art.1, Art.4, Art.5, Art.20	4

Figuur 2.10 Frequentieverdeling-tabel factoren en gevonden artikelen

De waargenomen voordelen van een data warehouse worden regelmatig genoemd in de inleidende woorden van een artikel: Een data warehouse ondersteunt business processen en besluitvorming. Ook technische factoren worden relatief vaak genoemd, omdat een data warehouse zich op dat vlak onderscheidt van gewone IT-systemen en daarmee interessant onderwerp is voor onderzoek en discussie. De Economische- en planningsfactoren scoren relatief laag ten opzichte van de overige factoren. Dit komt omdat er weinig onderzoek is gedaan naar operationele data warehouses en de beheerkosten ervan. Dit in tegenstelling tot data warehouse projecten en -implementaties.

Conclusie: De meetwaarden en succesfactoren uit het samengestelde model van paragraaf 2.2.4 worden bevestigd door de literatuur van de afgelopen 10 jaar.

2.4 Nieuwe meetwaarden

In de lijst met relevante artikelen en onderzoeken van paragraaf 8.3 vallen een paar zaken op, die regelmatig genoemd worden, zonder dat ze matchen met een bestaande meetwaarde. Hieronder worden deze verder uitgewerkt.

2.4.1 Agile development

Een Agile-werkwijze voor data warehouse development was in 2017 nog niet eerder onderzocht (Batra, 2017). De conclusie in dit artikel is dat de technologie-succesfactor (onder andere kundig IT-personeel, goede tools) sterk gerelateerd is aan een succesvolle agile-werkwijze, zoals Scrum. In een later artikel noemt Batra (2017) Agile-development in combinatie met klassiek projectmanagement als de ideale combinatie om een succesvol data warehouse neer te zetten.

Ook Watson et al. (2004) noemt in het artikel over governance, een incrementele werkwijze om het data warehouse stap-voor-stap te verbeteren. Het gebrek aan flexibiliteit (trage IT-organisatie) wordt genoemd door Huang et al. (2007) en Ramamurthy et al. (2007) als valkuil voor data warehouse succes.

Conclusie: de Agile-werkwijze moet genoemd worden in de meetwaarden van het model

2.4.2 Data Management

Data Management (DM) of Enterprise Data Management (EDM) is opgekomen in de jaren 90, tegelijk met de opkomst van data warehousing. Het omvat vele begrippen die toegepast worden op het hele data-domein. Tegenwoordig is Data Management onlosmakelijk verbonden met de internationale non-profit organisatie: DAMA International (DAMA, 2020). DAMA heeft in 2010 een boek uitgegeven (Mosley, 2010) en een model gelanceerd die inmiddels wereldwijd als de-facto standaard geldt. Het DM-model van DAMA ziet er als volgt uit:



Figuur 2.11 Data Management Body of Knowledge van DAMA

Data Management wordt genoemd in het artikel van Alhyasat (2013). Hier wordt gesproken over de nut en noodzaak van Data Management op individueel en corporate niveau. Ook het artikel over de data warehouse processen (Sen A. R., 2012) wordt Data Quality Management, Metadata Management en Data Governance genoemd om een hoog volwassenheidsniveau te halen. Hwang et al. (2008) vermelden in hun onderzoek dat er nog geen rekening is gehouden met Data Management. Thorne (2000) komt tot het inzicht dat beheer van datakwaliteit, data security en een gemeenschappelijke taal spreken (metadata) het meest belangrijk is om een succesvol data warehouse neer te zetten. Watson et al. (2004) benoemt het belang van afstemming van een data warehouse architectuur met Data Governance. Tot slot benoemt Al-Debei (2011) het hebben van een enterprise data model en het beheren van security en privacy als succesfactor voor een data warehouse.

Conclusie: Data Management moet genoemd worden in de meetwaarden van het model

2.4.3 Beheer

Een data warehouse is nooit klaar. In hun artikel geeft Hwang et al. (2008) aan dat er meer onderzoek nodig is naar beheer en groei van een data warehouse. Dat kan leiden tot een beheerfactor: "Easy to manage" (Hwang, 2008, p. 55). Ook het artikel over data warehouse processen (Sen A. R., 2012) gaat uitgebreid in op de stadia van volwassenheid van een data warehouse. Dit wordt bereikt door zaken als: Configuration management, Service Level Agreements (SLA's), Recovery management, Change management, Issue tracking, Problem management, Documentatie en meer van dit soort beheeronderwerpen. Ook in het onderzoek naar data warehouse adoptie (Ramamurthy, 2007) worden factoren genoemd als flexibiliteit en beheer. In het artikel over data warehouse governance (Watson, 2004) wordt vermeld dat het beheren van datakwaliteit een belangrijk aspect is. Huang et al. (2007) ziet dingen fout gaan omdat een data warehouse voortdurend groeit en onvoldoende schaalbaar is, wat tot problemen in het beheer en performance leidt.

Een aantal bestaande meetwaarden in het samengestelde model hebben een sterk operationeel karakter:

- De meetwaarden over schaalbaarheid, standaardisatie & documentatie als onderdeel van de factor systeemkwaliteit

- De meetwaarden over SLA's, Data Management en continue verbeteringen (Continuous Improvement - CI) als onderdeel van de technische factoren
- De meetwaarden over de registratie van verbeterpunten en het uitvoeren van wijzigingen (Changes) op het datawarehouse als planningsfactoren

Conclusie: Bovenstaande geeft voldoende aanleiding om een nieuwe succesfactor te benoemen: Beheer of beheerfactor. Bovenstaande meetwaarden zullen hiervan onderdeel uitmaken en de planningsfactor heeft dan geen meetwaarden meer en zal daarmee verdwijnen.

2.5 Conclusie Literatuurstudie

In paragraaf 2.2 is deelvraag 1a behandeld, over welke bestaande modellen bruikbaar zijn om de succesfactoren van een data warehouse in kaart te brengen. Er zijn 3 bruikbare modellen gevonden met hun meetwaarden, die in paragraaf 2.2.4 zijn geïntegreerd tot één samengesteld succesmodel. In paragraaf 2.3 zijn alle succesfactoren bevestigd door middel van een literatuurstudie van relevante artikelen van de afgelopen 10 jaar.

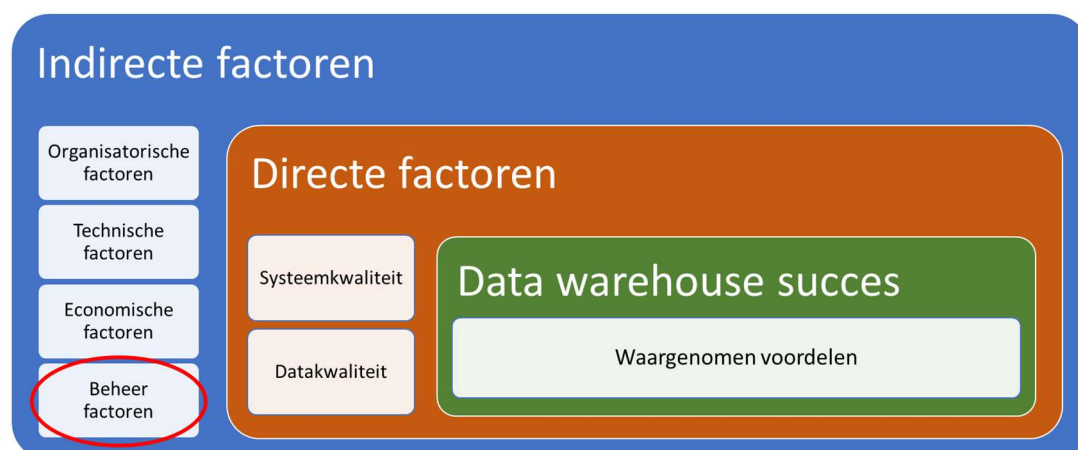
Deelvraag 2a, over welke nieuwe succesfactoren er de laatste jaren zijn bijgekomen, is in paragraaf 2.4 behandeld. In artikelen van de laatste 10 jaar worden een drietal onderwerpen vaak genoemd in combinatie met data warehousing. Deze zijn: Een Agile werkwijze om het bestaande data warehouse stap-voor-stap te verbeteren. Data Management, een vakgebied (en verzamelbegrip) in opkomst en waar het data warehouse onderdeel van uitmaakt. Dit leidt tot de nieuwe factor beheer, die ook de vorige twee begrippen omvat. Beheer als succesfactor ontbrak tot nu toe, omdat vrijwel alle bestaande modellen uitgaan van data warehouse-projecten en -implementaties.

Factor / Categorie	Nr	Meetwaarde / Beschrijving
Waargenomen voordelen	V1	Ondersteunen primaire, secundaire en innovatieve processen
	V2	Ondersteunen van strategische doelstelling
	V3	Verbeterde individuele productiviteit
	V4	Verbeterde besluitvorming op strategisch en tactisch niveau
	V5	Verbeterde besluitvorming op operationeel niveau
	V6	Ondersteunen van analyse-doeleinden
Systeemkwaliteit	S1	Makkelijk in het gebruik en gebruiksvriendelijk
	S2	Eenvoudig en efficiënt
	S5	Security en autorisatie is goed geregeld
	S6	Integratie van meerdere bronsystemen
	S7	Goede beschikbaarheid
Datakwaliteit	D1	Hoogwaardige datakwaliteit
	D2	Voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens
	D3	Toereikende actualiteit van data (laadfrequentie en verversingsgraad)
	D4	Toegang tot geïntegreerde en consistente data
Organisatorische factor	O1	Doelstellingen zijn bekend en voldoen aan de business behoeften
	O2	Management vanuit de business en IT is betrokken en ondersteunt
	O3	Business (eind)gebruikers zijn actief betrokken
	O4	Goede communicatie tussen IT en business
	O5	Resources zijn beschikbaar gesteld
	O6	Er is een visie en roadmap
Technische factor	T3	Goede ontwikkelomgeving
	T4	Kundig IT-personeel
	T6	Veel data en bronsystemen, waardoor technisch complex
Economische factor	E1	Voldoende funding
	E2	Voordelen zijn bekend en worden financieel gemeten
	E3	Grootte organisatie
Beheer	B1	Verbeterpunten in de vorm van een backlog zijn bekend en geregistreerd
	B2	Continue verbeteringen door Agile werkwijze
	B3	Wijzigingen worden goed uitgevoerd en gemanaged
	B4	Goede responsetijd door schaalbaarheid
	B5	Gestandaardiseerd, gedocumenteerd en onderhoudbaarheid
	B6	Goede datakwaliteit aanleverende bronsystemen
	B7	Data Management is ingericht

Figuur 2.12 Dataverzameltabel na literatuurstudie

In het rood staan de nieuwe meetwaarden genoemd.

Het nieuwe samenstelde data warehouse succesmodel ziet er als volgt uit:



Figuur 2.13 Nieuw samengesteld model succesfactoren data warehouses

3 Case Study

Nu de theoretische basis is neergezet, wordt in dit hoofdstuk gekeken naar de methodische aanpak van het onderzoek. Het samengestelde model, zoals beschreven in paragraaf 2.5 is de theorie waarop dit onderzoek verder gaat. De focus is hierbij op de deelvragen 1b, 1c, 2b, 2c: In hoeverre gelden de succesfactoren voor grote financiële organisaties in Nederland met (langdurig) operationele data warehouses?

3.1 Onderzoeksmethode en -ontwerp

De onderzoeksfunctie bij dit onderzoek is: Verklaren (explain). Als uitgegaan wordt van een bestaande theorie, dan is de onderzoeksmethode: deductief en theory testing (Bryman, 2015). In het boek Business Research Methods geeft Bryman et al. aan dat het proces van deductie begint bij een theorie, waarna de datacollectie volgt. Vervolgens wordt de data geanalyseerd en de uitkomsten vergeleken met de bestaande theorie. Dit kan leiden tot een bevestiging of een revisie van de bestaande theorie.

In de onderzoeksvraag zitten 3 aspecten (variaties) die vergeleken zijn met de bestaande theorie:

1. **Locatie:** De theorie komt uit de Verenigde Staten, maar is deze ook toepasbaar in Nederland?
2. **Tijd / Actualiteit:** De theorie komt uit de jaren nul van deze eeuw, maar is deze nu ook nog geldig?
3. **Onderwerp:** De theorie gaat over succesfactoren bij data warehouse implementaties, maar gelden deze ook bij bestaande operationele data warehouses?

Het onderzoeksontwerp (research design) is kwalitatieve case study, zoals Robert K. Yin (2013) dit beschreven heeft. Dit betekent dat het model getoetst is, aan de hand van enquêtes en interviews (cases) met relevante functionarissen en organisaties. In zijn boek Case Study Research beschrijft Yin (2013) drie belangrijke voorwaarden die de keuze voor een passende onderzoeksmethode bepalen. Ten eerste is het type onderzoeksvraag van belang, waarbij een verklarende vraag ('hoe' of 'waarom') goed past bij een case study aanpak, die vaak de diepte opzoekt. Ten tweede is de behoefte aan controle over, en beïnvloeding van de te onderzoeken situatie van belang. Is deze behoefte aanwezig, dan wordt vaak voor een experiment gekozen. Ten derde is van belang of het onderzoek zich richt op actuele en hedendaagse gebeurtenissen.

In dit onderzoek gaat het om een verklarende vraag: 'Wat zijn de factoren?'. Er is geen behoefte aan het beïnvloeden van de te onderzoeken situaties en er is sprake van onderzoek naar de actuele hedendaagse situatie. Deze voorwaarden komen overeen met de keuze voor een case study als passend onderzoeksontwerp.

In haar artikel over Case Study Survey, gaat Chmiliar (2012) dieper in op het onderzoeksontwerp, door specifiek de onderwerpen: enquête en interview te benoemen. Een survey in de vorm van een enquête, is met name geschikt als het gaat om een kleine steekproef om een gehele populatie te karakteriseren. De individuele enquêtes of interviews geven een totaalbeeld van de gehele populatie. Dit is dus goed van toepassing op dit onderzoek, waarbij via een steekproef de hele populatie van alle grote financiële dienstverleners in Nederland in kaart gebracht is.

In het verleden is de combinatie van Theory Testing en Case Studies vaak onderwerp van discussie geweest. In een onderzoek van Lokke (2014) wordt uitgelegd dat een Case Study juist goed kan bijdragen aan het verklarende deel van de bestaande theorie. Ook de grenzen van de theorie worden hiermee in kaart gebracht.

Volgens Paré (2004) is een Case Study een goed mechanisme als er sprake is van IT-omgevingen waar een grote dynamiek en veranderende condities gelden. Met name in de hoek van systeemontwikkeling, maar ook IT-management, kan een Case Study helpen bij het onderzoeken van IT-impact op organisaties (Paré, 2004). Het onderzoek van succesfactoren op een data warehouse als IT-component sluit hier prima op aan.

3.2 Onderzoekspopulatie

In Nederland hebben vooral de grote bedrijven geïnvesteerd in data warehouse oplossingen. Dit onderzoek richt zich op grote financiële instellingen, omdat die al langer de nut en noodzaak van data warehouses en Business Intelligence inzien (Ramamurthy, 2007), waardoor ze meer dan gemiddeld ervaring hebben met het beheer van (langdurig operationele) data warehouses.

Volgens het CBS zijn er 25 financiële instellingen van 1.000 medewerkers of meer (CBS, 2020) in het 4^e kwartaal van 2019:



Figuur 3.1 CBS-informatie over bedrijfsgrootte in Nederland

Dit onderzoek is representatief voor de 25 grote financiële dienstverleners in Nederland met meer dan 1000 werknemers, bij een steekproef van 8 organisaties, bij een betrouwbaarheidsniveau van 80% en een foutmarge van 20% (Steekproefcalculator, 2020).

3.3 Enquête en interview

In deze paragraaf wordt beschreven hoe de vragenlijst als basis voor de enquête en interviews tot stand is gekomen. De enquête is online aangeboden met behulp van een enquête-tool Easion Survey (Easion.nl, 2020) en bestaat uit algemene vragen, gesloten vragen over de succesfactoren en twee open vragen over data warehouse ontwikkelingen en alternatieven.

3.3.1 Participanten

Een vragenlijst in enquêtevorm is verstuurd naar financiële organisaties in Nederland die (langdurig) ervaring hebben met bestaande operationele data warehouse omgevingen. De deelnemers hieraan zijn gevonden in het directe netwerk van de onderzoeker en begeleider. Daarnaast is de enquête online gezet voor een selecte club van data warehouse experts via

het social media platform Slack. Dit levert voldoende bruikbare vragenlijsten op van tenminste 8 verschillende banken en verzekeraars, die voldoen aan de gestelde criteria.

3.3.2 Vragenlijst en enquête

De vragenlijst voor de enquête is een vertaling van de meetwaarden uit de dataverzameltabel van paragraaf 2.5 naar concrete hedendaagse vragen voor de financiële dienstverleners in Nederland. In bijlage 8.4 wordt getoond hoe de vertaling heeft plaatsgevonden. Alle 34 meetwaarden, behorend bij de 7 directe en indirecte succesfactoren uit het samengestelde model, zijn omgezet naar vragen voor de online enquête in de online enquête-tool Easion (Easion.nl, 2020). Hierdoor wordt het volledige model getoetst bij de betreffende organisaties. De volgorde van de vraagstelling komt grotendeels overeen met de volgorde van de dataverzameltabel. Door voortschrijdend inzicht tijdens de literatuurstudie, is het model wat verder doorontwikkeld, nádat de enquête online is gezet. Dit heeft geen invloed op de vraagstelling en uitkomst van de enquête.

Naast de gesloten vragen uit het samengestelde succesmodel, worden er ook een aantal open vragen gesteld. Het eerste deel van de open vragen gaat over de indicatoren die nodig zijn om te beoordelen of de participant in zijn/haar organisatie voldoet aan de gestelde criteria van dit onderzoek. De laatste twee open vragen gaan over deelvraag 3a en 3b: Welke moderne alternatieven of ontwikkelingen gelden er voor data warehouses. Dit onderdeel zal tijdens de interviews ook aan bod komen en wordt hieronder verder uitgewerkt.

3.3.3 Interviews

Naast de direct aangeschreven deelnemers voor een interview, kan in de online enquête aangegeven worden of er een interview mag volgen over de gestelde antwoorden. De interviews worden gehouden bij grote Nederlandse financiële organisaties die één of meerdere data warehouse omgevingen operationeel hebben. De interviews zijn semigestructureerd, wat betekent dat veel vragen, die ook in de enquête staan, besproken worden om meer achtergrondinformatie te vergaren. Daarnaast wordt de nadruk gelegd op de open vragen, waarin de deelnemers de kans krijgen om zaken in hun organisaties nader toe te lichten. Voorbeelden van de functies van de geïnterviewden, zijn: Data warehouse manager, Business Intelligence manager, Informatiemanager, Solution architect, Data architect, Data warehouse beheerder of (senior) data warehouse ontwikkelaars.

De uitkomsten van de interviews, alsmede de beantwoording van de open vragen, geven inzicht in de ontwikkelingen op het data-domein in een organisatie.

3.3.4 Analyse

De resultaten van de enquêtes en de interviews worden kwalitatief geanalyseerd en gecategoriseerd.

Analyse algemene vragen uit de enquête om te bepalen of de respondenten tot de doelgroep behoren:

- Groeperen, filteren en rapporteren door middel van Excel

Analyse gesloten enquêtevragen ter beantwoording van de deelvragen 1b en 1c:

- Verzamelen van alle enquête-antwoorden in het online survey-tool Easion (Easion.nl, 2020);
- Filteren van volledig ingevulde enquêtes die voldoen aan de criteria van de doelgroep in Easion;
- Groeperen van antwoorden per meetwaarde / succesfactor in Easion;
- Rapport opmaken van de resultaten per deelnemer vanuit Easion;
- Rapport opmaken van de resultaten per case per meetwaarde vanuit Excel.

Analyse interviews en open enquêtevragen ter beantwoording van de deelvragen 2a, 2b, 2c, 3a en 3b:

- De interviews worden eerst getranscribeerd. Dit betekent dat de geluidsopnames van de interviews woordelijk uitgeschreven worden in een tekstbestand;
- In de QDA-tool (Qualitative Data Analysis) ATLAS.ti (ATLAS.ti, 2020) worden de teksten ingelezen;
- Ook de enquêteresultaten van de open vragen worden in ATLAS.ti ingelezen;
- De teksten worden in ATLAS.ti gecodeerd en gecategoriseerd;
- De codes zijn deels te koppelen aan de vragen uit de enquête, en deels zijn het nieuwe categorieën of inzichten, wat kan leiden tot nieuwe succesfactoren (deelvraag 2) of ontwikkelingen (deelvraag 3);
- Rapport opmaken van de frequentieverdeling, wat aangeeft hoe vaak de codes genoemd worden, vanuit ATLAS.ti;
- Rapport opmaken van de code-document verhouding, die aangeeft bij welke cases de codes gebruikt zijn, vanuit ATLAS.ti;
- Rapport opmaken van de co-occurrences, die aangeven welke codes aan elkaar gerelateerd zijn, vanuit ATLAS.ti.

3.4 Ethische verantwoording

Deelnemers aan de enquête en/of interview hebben zelf de keuze om de vragen anoniem te beantwoorden. Ook kunnen ze ervoor kiezen om de naam van hun organisatie niet te vermelden. De privacy van de deelnemers en hun organisatie is gewaarborgd, want de individueel ingevulde gegevens van de vragenlijsten worden niet openbaar gemaakt en daarmee is persoonlijk identificeerbare informatie nooit traceerbaar. De individueel gegeven antwoorden zullen confidentieel worden behandeld. De informatie die hieruit voortvloeit wordt puur gebruikt ter onderbouwing van het onderzoek en zal buiten dit onderzoek niet gepubliceerd worden.

Omdat de meeste participanten aangegeven hebben de naam van hun organisatie niet te willen publiceren, is ervoor gekozen om geen enkele naam te vermelden, maar alleen cases te noemen: Case1, Case2, etc. Dit is in lijn met wat Bryman et al. Beschrijven in hun boek (2015).

De belangrijkste stimulans voor deelname aan dit onderzoek is dat de resultaten van dit onderzoek, indien gewenst, teruggekoppeld kunnen worden. Ook hierbij is privacy gewaarborgd, zoals hierboven beschreven.

4 Resultaten

In dit hoofdstuk worden de resultaten van de enquête en de interviews kwalitatief gepresenteerd per succesfactor. In paragraaf 4.1 wordt er een beeld geschetst van de respondenten met de algemene kenmerken van hun organisaties en of deze voldoende representatief zijn voor het beantwoorden van de deelvragen. In paragraaf 4.2 worden de resultaten van de enquête uitgewerkt per succesfactor. In paragraaf 4.3 wordt gekeken naar de resultaten van de open vragen uit de enquête en de interviews, om nieuwe ontwikkelingen bij grote Nederlandse organisaties te ontdekken.

4.1 Respondenten

Er zijn 82 potentiële deelnemers benaderd, waarvan 22 direct via het netwerk van de onderzoekers en 60 deelnemers via een vereniging van samenwerkende specialisten op het gebied van BI en data warehousing. Dit heeft geresulteerd in 18 volledig ingevulde enquêtes. Van de 18 respondenten zijn er 3 op locatie geïnterviewd, om meer achtergrondinformatie te verzamelen. Daarnaast heeft 1 interview plaatsgevonden met een respondent die de enquête gedeeltelijk had ingevuld. In totaal dus 19 respondenten.

De kenmerken van deze respondenten zijn:

- De 19 respondenten vertegenwoordigden 11 verschillende organisaties in Nederland;
- Allen werkten voor grote organisaties in Nederland variërend van 1.500 tot 40.000 werknemers.

#	Sector	# Werknemers	Functie	# Data warehouses	Grootte	Leeftijd	Doelgroep ja/nee
1	Case 1 Verzekeringen	3.500-4.000	DWH ontwerper	8-10	8-16Tb	8-10 jaar	Ja
2	Verzekeringen	3.500-4.001	Risk expert	8-11	8-16Tb	8-10 jaar	Ja
3	Verzekeringen	3.500-4.002	DWH ontwikkelaar	8-12	8-16Tb	8-10 jaar	Ja
4	Case Staffing	3.000	Project Manager DWH	2	5Tb	10 jaar	Nee
5	Staffing	3.000	DWH specialist	2	5Tb	10 jaar	Nee
6	Case 2 Verzekeringen	1.600	BI consultant	3	2Tb	8 jaar	Ja
7	Case 3 Verzekeringen	3.000	System Architect	4	2Tb	10 jaar	Ja
8	Case Provincie	1.500	Informatiemanager	7	20Gb	11 jaar	Nee
9	Case Ziekenhuis	12.000	Data Architect	4	5Tb	8 jaar	Nee
10	Case 4 Verzekeringen	3.450	Functioneel Beheer BI	10	2,5Tb	10 jaar	Ja
11	Case 5 Bank	40.000	Business Analyst	10-60	10Tb	2-5 jaar	Ja
12	Bank	40.000	DWH ontwikkelaar	10-60	10Tb	2-5 jaar	Ja
13	Case 6 Verzekeringen	15.000	Informatie-architect	5-8	6Tb	5-10 jaar	Ja
14	Verzekeringen	15.000	DWH-ontwikkelaar	5-8	6Tb	5-10 jaar	Ja
15	Case 7 Bank	20.000	IT Run Specialist	2-5	25-50Tb	10-15 jaar	Ja
16	Bank	20.000	Senior Business Analyst	2-5	25-50Tb	10-15 jaar	Ja
17	Bank	20.000	Product Owner	2-5	25-50Tb	10-15 jaar	Ja
18	Bank	20.000	Development Engineer	2-5	25-50Tb	10-15 jaar	Ja
19	Case 8 Bank	50.000	Data management analyst	100?	?	4 jaar	Ja

Figuur 4.1 Kenmerken respondenten en selectie doelgroep

Doelgroep

In de onderzoeksvraag wordt gekeken naar operationele data warehouses bij grote financiële organisaties in Nederland en daar voldoen 15 van de 19 ondervraagden aan. Deze 15 werken voor 8 verschillende grote Nederlandse financiële organisaties (cases). De andere 4 maken geen deel uit van een financiële dienstverlener, maar voldoen wel aan de overige criteria en zijn daarmee een waardevolle bijvangst in het onderzoek.

Overige kenmerken van de respondenten uit de doelgroep zijn:

- Gemiddeld waren er 8 data warehouses per organisatie actief, waarbij één uitschieter met een geschat aantal van 60 (of misschien wel 100, maar dit was onzeker);
- De gemiddelde leeftijd van de data warehouses is 8 jaar;
- De gemiddelde grootte van de data warehouses is 13 TB, alhoewel een aantal respondenten aangeven heeft dat het een schatting is.

Conclusie: De steekproef onder de bovenstaande doelgroep is **voldoende representatief** (zie criteria in paragraaf 3.2) om de onderzoeksvraag en de deelvragen te kunnen beantwoorden.

4.2 Resultaat per succesfactor

In de volgende paragrafen worden de resultaten uit de enquête per succesfactor geanalyseerd en gepresenteerd. In bijlage 8.5 staat het totale resultaat van de enquête in Excel uitgewerkt.

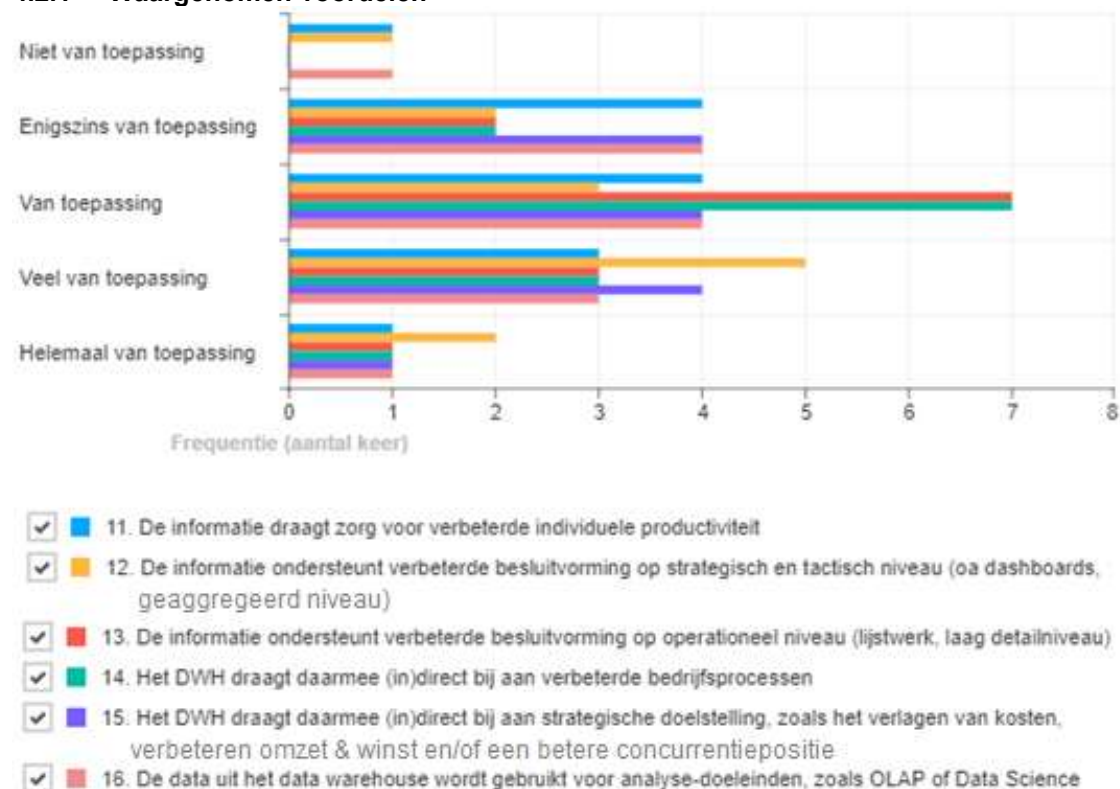
De kwalitatieve analyse is tweeledig:

1. De grafieken uit Easion (Easion.nl, 2020) tonen het aantal antwoorden (frequentie) dat gegeven is per vraag per succesfactor. Dit geeft een beeld van de meest voorkomende antwoorden en de spreiding ervan;
2. Daarnaast wordt bekeken wat het resultaat is per case per meetwaarde in de dataverzameltabel. Op deze manier kan een onderzoeker aantonen wat er gezegd is en waarop de kwalitatieve resultaten onderbouwd zijn (Bloomberg, 2008). Een score van: “Van toepassing”, “Veel van toepassing” en “Helemaal van toepassing” is groen gekleurd en telt mee in het percentage in de laatste kolom. Een score van “Niet van toepassing” en “Enigszins van toepassing” is bruin gekleurd en telt niet mee in het percentage. Per case wordt er afgerond naar beneden, dat wil zeggen dat 2 antwoorden: “Enigszins van toepassing” en “Van toepassing”, resulteert in een “Enigszins van toepassing”-uitslag.

Het percentage is geen kwantitatieve beoordeling, maar geeft een kwalitatieve indruk van de mate van toepasbaarheid van de succesfactoren bij de onderzoeksgroep.

Met bovenstaande analyses van de enquêteresultaten per succesfactor wordt gekeken of de bestaande en nieuwe succesfactoren en hun meetwaarden gelden voor grote financiële dienstverleners in Nederland. Daarmee worden in hoofdstuk 5 de deelvragen 1b, 1c, 2b en 2c beantwoord.

4.2.1 Waargenomen voordelen



Figuur 4.2 Enquêteresultaten waargenomen voordelen. Bron: Easion Survey

De vragen 11 t/m 16 gaan over de voordelen die een data warehouse biedt voor de individuele werknemers en de gehele organisatie. Opvallend is dat, op één na alle vragen een neutrale score “Van toepassing” scoren, zodat algemeen gesteld kan worden dat de waargenomen voordelen van een data warehouse gelden voor grote Nederlandse organisaties. Het ondersteunen van verbeterde besluitvorming op strategisch en tactisch niveau (vraag 12) was overwegend veel van toepassing, wat betekent dat het data warehouse vaak dient als basis voor management-informatie, bijvoorbeeld in de vorm van dashboards. Wel toont figuur 4.2 aan dat de antwoorden sterk verschilden hierbij. Dit gold ook voor de meeste andere antwoorden bij deze factor van waargenomen voordelen. Deze reden van deze spreiding wordt later toegelicht in paragraaf 5.1

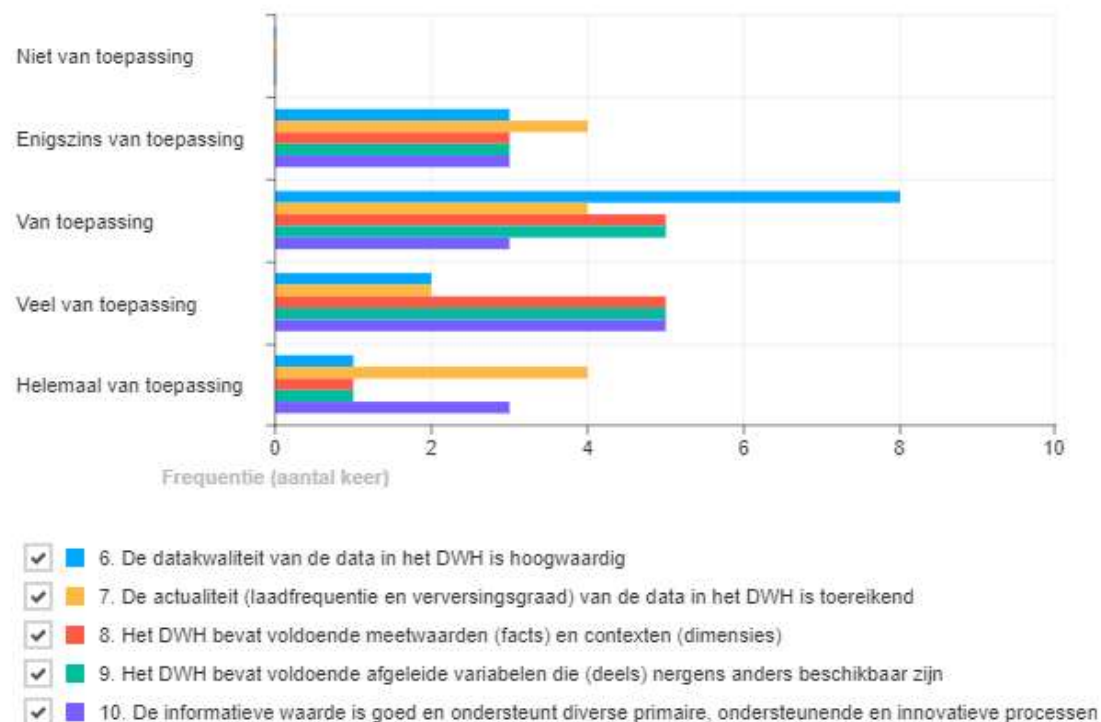
In figuur 4.3 worden de meetwaarden voor de 8 cases in de dataverzameltabel getoond:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V15 + V14	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	87,50%
V11	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	62,50%
V12	Veel van toepassing	Veel van toepassing	Niet van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	87,50%
V13	Van toepassing	Veel van toepassing	Van toepassing	Helemaal van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Veel van toepassing	87,50%
V16	Van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	75,00%

Figuur 4.3 Percentage cases van toepassing per vraag - Factor waargenomen voordelen

De hoge percentages duiden erop dat de waargenomen voordelen van een data warehouse gelden voor vrijwel alle financiële dienstverleners. Alleen case 3 is een uitzondering hierop, maar dit is te verklaren, doordat er een nieuw centraal Data Lake ontwikkeld wordt, die ook de functies van het bestaande data warehouse moet gaan overnemen.

4.2.2 Datakwaliteit



Figuur 4.4 Enquête resultaten datakwaliteit. Bron: Easion Survey

De vragen 6 t/m 10 gaan over datakwaliteit. Hier waren de antwoorden meer eenduidig. De hogere scores bij vraag 9 en 10 geven aan dat de operationele data warehouses bij financiële

dienstverleners in Nederland voldoende meetwaarden, contexten en afgeleiden gegevens bevatten met een goede informatieve waarde en daarmee toegevoegde waarde levert.

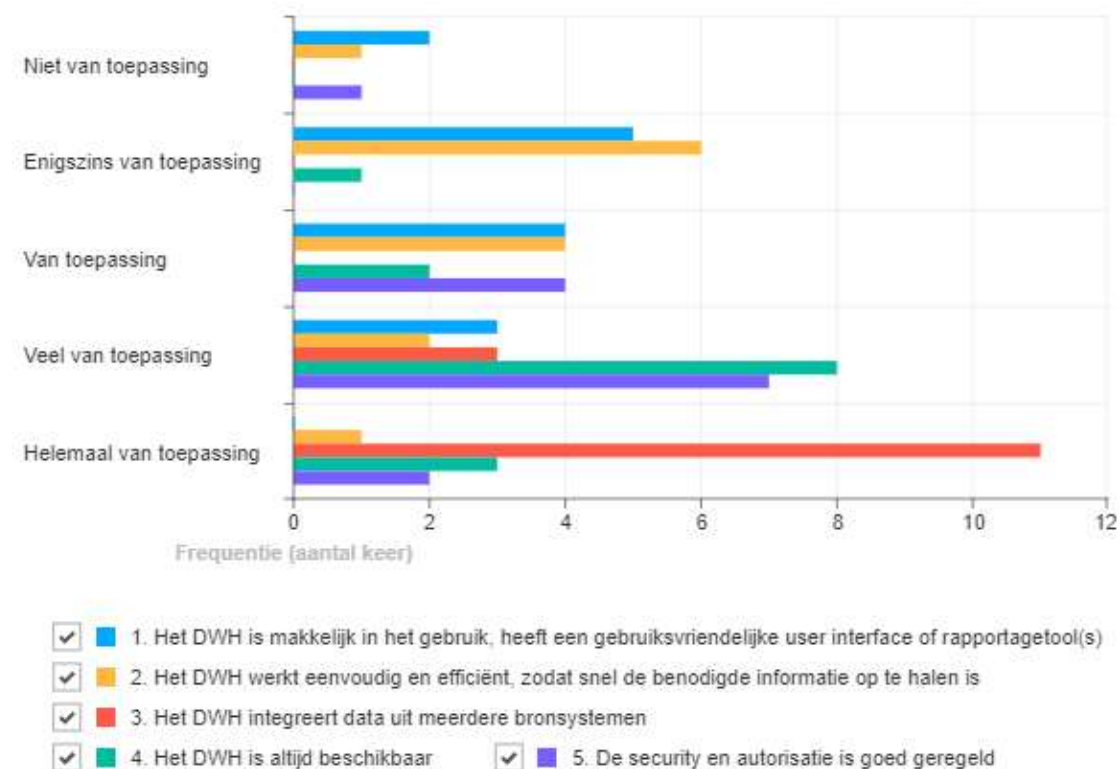
De meetwaarden per case geven het volgende beeld:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V10	Van toepassing	Helemaal van toepassing	Enigszins van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	87,50%
V06	Van toepassing	Van toepassing	Veel van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	87,50%
V08 + V09	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	75,00%
V07	Van toepassing	Helemaal van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	75,00%
V08 + V09	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	75,00%

Figuur 4.5 Percentage cases van toepassing per meetwaarde - Factor datakwaliteit

Ook hier hoge percentages, dus datakwaliteit is van toepassing bij de operationele data warehouses bij de onderzochte organisaties. Case 1, 3 en 6 zijn alle drie verzekeraars en hebben de datakwaliteit minder op orde. Dit komt grotendeels door de verzuilde organisatiestructuur, waarbij de business units vrij autonoom van elkaar opereren, waardoor er een versnipperd datalandschap ontstaat met meerdere data warehouses. Dit komt de datakwaliteit niet ten goede.

4.2.3 Systeemkwaliteit



Figuur 4.6 Enquête resultaten systeemkwaliteit. Bron: Easion Survey

De vragen 1 t/m 5 gaan over systeemkwaliteit. Opvallende score is dat alle data warehouses data integreren uit meerdere systemen (vraag 3). Eén van de oorspronkelijke theoretische doelen van een data warehouse wordt hiermee in de praktijk bewezen. Ook de beschikbaarheid en security zijn goed geregeld, met een hoge score bij vraag 4 en 5. Daarmee is deze succesfactor sterk aanwezig bij de grote Nederlandse financiële dienstverleners.

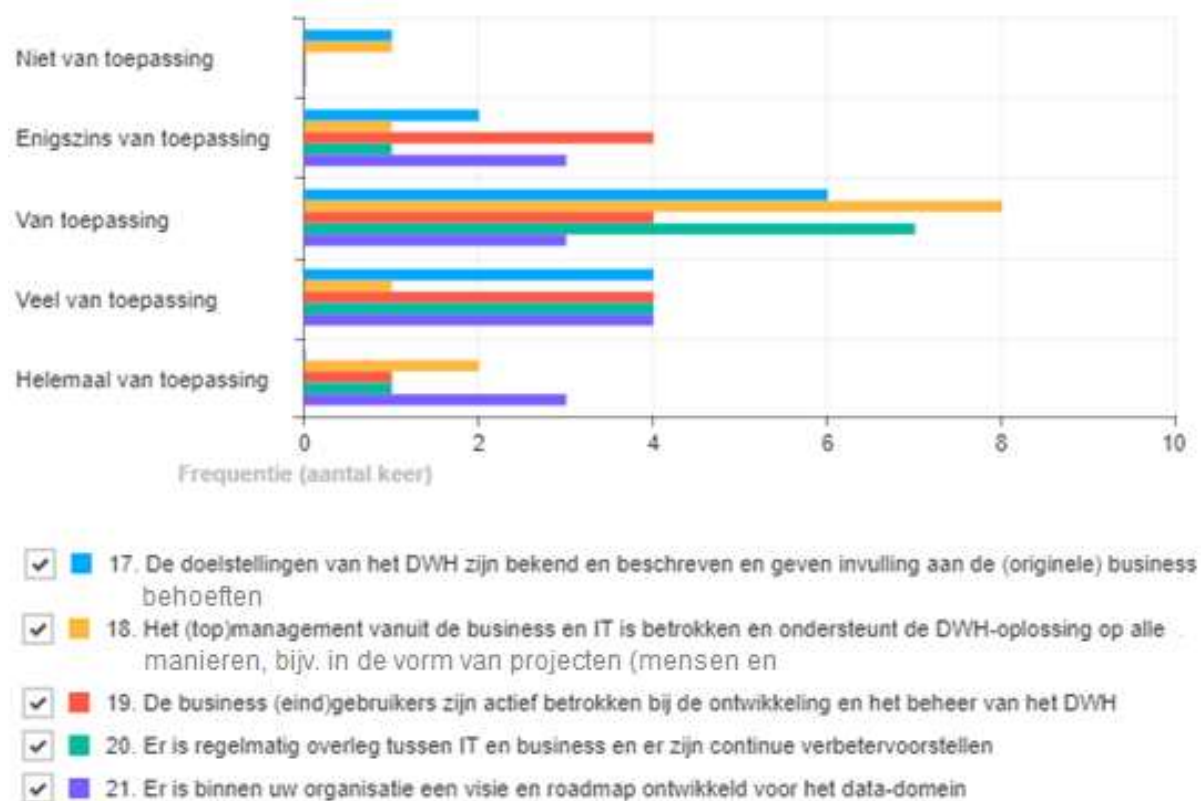
Dit blijkt ook uit de meetwaarden per case, zoals uit onderstaand figuur 4.7 blijkt:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V01	Enigszins van toepassing	Veel van toepassing	Niet van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	62,50%
V02	Veel van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	50,00%
V05	Veel van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	100,00%
V03	Veel van toepassing	Veel van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	100,00%
V04	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	100,00%

Figuur 4.7 Percentage cases van toepassing per meetwaarde - Factor systeemkwaliteit

Drie hoge scores van 100% geven aan dat alle financiële organisaties hun systeemkwaliteit goed op orde hebben. Wel zijn de meningen sterk verdeeld als het gaat om de eenvoud en efficiëntie. Dit is te verklaren door te kijken naar de functies van de deelnemers. De functies die dicht bij de business staan, zijn minder enthousiast over de eenvoud en efficiëntie van het data warehouse, dan de IT-ers.

4.2.4 Organisatorische factoren



Figuur 4.8 Enquête resultaten organisatorische factoren. Bron: Easion Survey

De vragen 17 t/m 21 gaan over de organisatie en de betrokkenheid ervan. Met een overwegend: "Van toepassing", geldt dat het management, de business en IT voldoende betrokken zijn om deze succesfactor te kunnen bevestigen. Net als in de interviews bevestigd wordt, blijkt dat er vaak een visie en roadmap ontwikkeld is of wordt. Enerzijds omdat de financiële organisaties in Nederland data hoog op de agenda hebben staan, maar anderzijds ook om het huidige datalandschap te vervangen voor een modernere omgeving. Dit laatste wordt verder besproken in de ontwikkelingen in paragraaf 4.3.

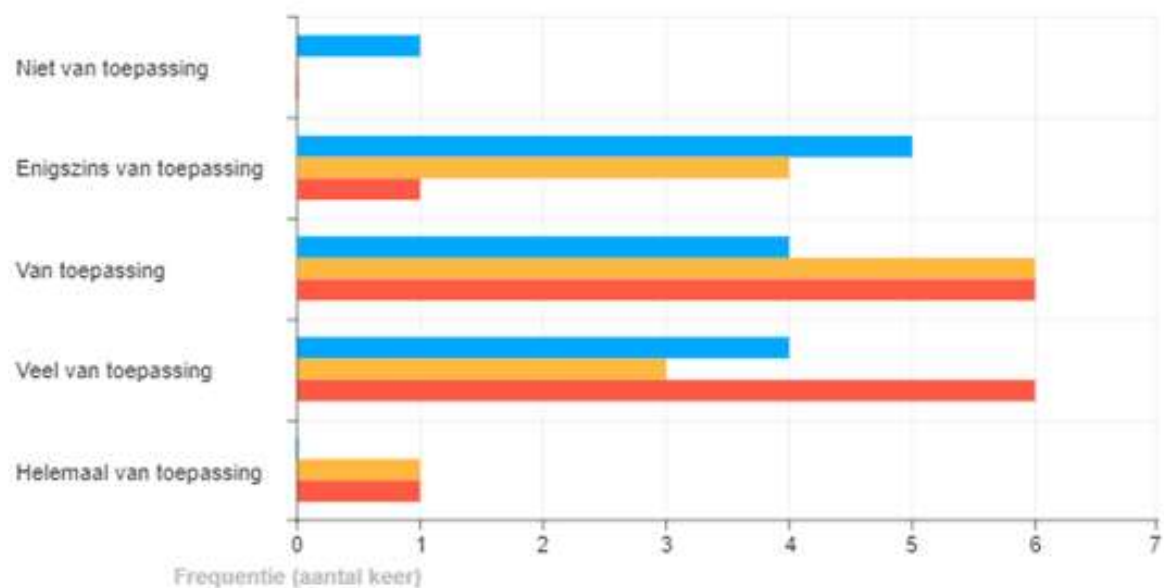
Figuur 4.9 toont de score per case per meetwaarden:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V17	Van toepassing	Veel van toepassing	Niet van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	87,50%
V18	Van toepassing	Enigszins van toepassing	Niet van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	75,00%
V19	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	75,00%
V20	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	100,00%
V21	Enigszins van toepassing	Van toepassing	Helemaal van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	87,50%

Figuur 4.9 Percentage cases van toepassing per meetwaarde – Organisatorisch factoren

Opvallend hierbij is de 100% score bij de samenwerking en communicatie tussen IT en de business. Blijkbaar is dat prima geregeld bij alle financiële dienstverleners in Nederland. Ook de overige percentages zijn hoog, waardoor vast te stellen is dat de organisatorische factoren goed op orde zijn in Nederland.

4.2.5 Technische factoren



- ☒ 22. De bronsystemen leveren data van voldoende kwaliteit: juist, volledig en tijdig. Dit is contractueel vastgelegd
- ☒ 23. De technologie van het DWH is toereikend. De tools en databases zijn actueel en worden ondersteund door de leveranciers
- ☒ 24. De kennis en kunde van het IT-personeel is toereikend. Er wordt gebruik gemaakt van externe consultants zodra dat nodig is

Figuur 4.10 Enquêteresultaten technische factoren. Bron: Easion Survey

De vragen 22 t/m 24 gaan over de technische factoren. Hier geldt een hoge score voor kennis en kunde van het IT-personeel. Dit komt niet alleen omdat veel respondenten uit de IT komen en zichzelf niet willen tekortdoen, maar ook omdat veel data warehouses in het verleden vanuit IT “gepushed” werden richting de business.

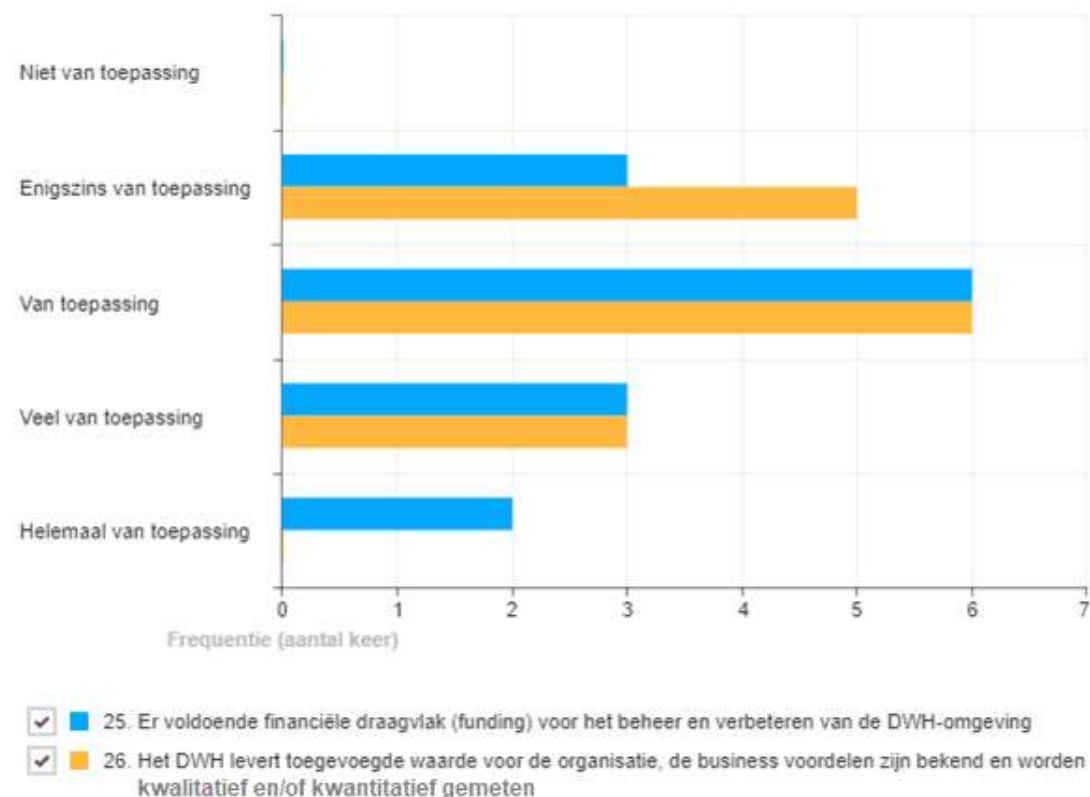
De meetwaarden per case geven het volgende beeld:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V23	Van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	100,00%
V24	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	100,00%
V22 + V32	Enigszins van toepassing	Enigszins van toepassing	Niet van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	50,00%

Figuur 4.11 Percentage cases van toepassing per meetwaarde – Technische factoren

Hier valt de relatief lage score op van 50% van de aanleverende bronsystemen. Vanuit de interviews is gebleken dat het niet alleen om de kwaliteit van de bronsystemen ging, maar ook aan het ontbreken van een contractuele vastlegging.

4.2.6 Economische factoren



Figuur 4.12 Enquêteresultaten economische factoren. Bron: Easion Survey

De twee vragen over economische factoren zijn beiden van toepassing. De funding lijkt prima op orde. Vanuit de interviews en de literatuurstudie is gebleken, dat dit grotendeels komt omdat een Agile-werkwijze is omarmt. Hierbij wordt niet gekeken naar de kosten per wijziging of uitbreiding op het data warehouse, maar wordt een compleet Scrum- of Devops-team gefinancierd voor een bepaalde periode. Het betekent dat het data warehouse nog steeds relevant geacht wordt en voordelen biedt die soms ook gemeten worden.

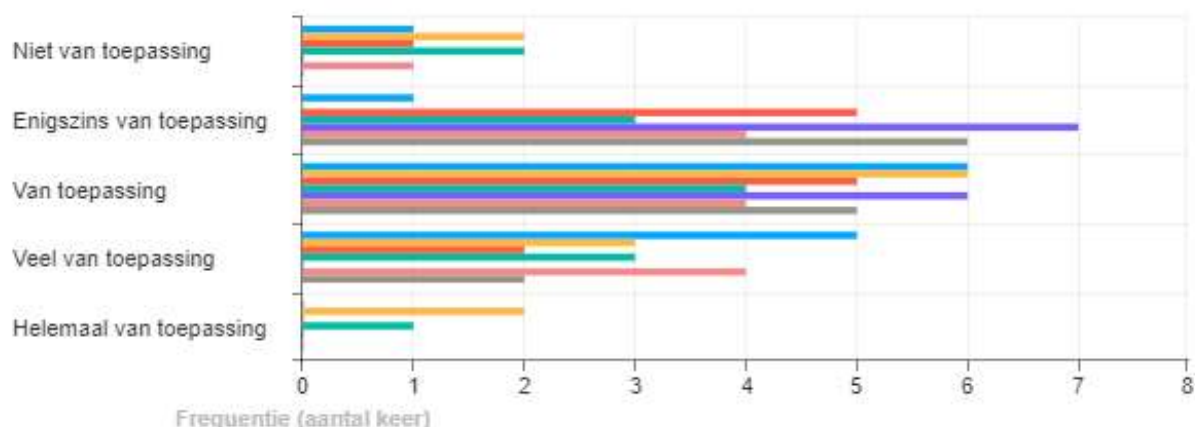
Figuur 4.13 toont de score per case per meetwaarden:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V25	Van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	75,00%
V26	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	50,00%

Figuur 4.13 Percentage cases van toepassing per meetwaarde – Economische factoren

De vrij lage score van 50% is besproken tijdens de interviews. De toegevoegde waarde van een data warehouse wordt niet onderschat, maar er worden geen metingen gedaan om dit financieel te onderbouwen, bijvoorbeeld in de vorm van een business case.

4.2.7 Beheerfactoren



- ☒ 27. Er is een roadmap of backlog met eisen en wensen voor de (nabije) toekomst
- ☒ 28. Het DWH wordt continue verbeterd door een agile werkwijze of door regelmatig IT-projecten uit te voeren
- ☒ 29. De wijzigingen op het DWH worden telkens tijdig opgeleverd, binnen budget, scope en kwaliteitseisen
- ☒ 30. Het DWH is schaalbaar vwb de hardware en databases waarop het draait
- ☒ 31. Het DWH is gestandaardiseerd, gedocumenteerd en eenvoudig te onderhouden
- ☒ 32. De bronsystemen leveren data van voldoende kwaliteit: juist, volledig en tijdig. Dit is vastgelegd in een overeenkomst (SLA of DDA)
- ☒ 33. Data management is ingericht en er wordt gewerkt met tools rond metadata, master data en data profiling

Figuur 4.14 Enquêteresultaten beheerfactoren. Bron: Easion Survey

De beheer-factoren worden gevraagd bij de vragen 27 t/m 33. Ondanks een neutrale score van “Van toepassing” op alle onderdelen blijkt de spreiding aan antwoorden erg groot. Ook tijdens de interviews werd dit gemerkt. Waar de ene organisatie volop investeert in data warehousing en goede rapportage, is een andere organisatie heel ver met de opvolging van een data warehouse. Dit komt vooral tot uiting in de beheerfactoren: Wordt er nog veel tijd en geld gestoken in het beheer of investeert men liever in de moderne alternatieven?

Ook de resultaten per case geven dit beeld:

Vr Nr	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	% (N=8)
V27	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	75,00%
V28	Van toepassing	Helemaal van toepassing	Niet van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Helemaal van toepassing	75,00%
V29	Van toepassing	Veel van toepassing	Niet van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	62,50%
V30	Van toepassing	Enigszins van toepassing	Niet van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	62,50%
V31	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	25,00%
V22 + V32	Enigszins van toepassing	Enigszins van toepassing	Niet van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	50,00%
V33	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	37,50%

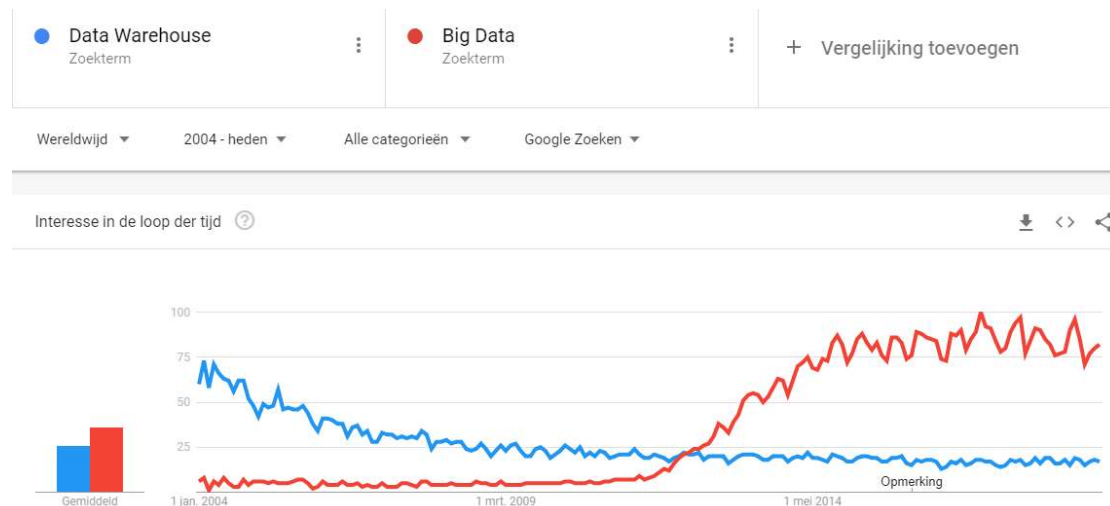
Figuur 4.15 Percentage cases van toepassing per meetwaarde – Beheerfactoren

In het algemeen kan gesteld worden dat de beheerfactoren de minste invloed hebben op het succes van een operationeel data warehouse bij de financiële dienstverleners in Nederland. De nieuwe gevonden meetwaarde rond Agile development is sterk van toepassing (75%). Ronduit slecht gesteld is het met de standaardisatie en documentatie van een data warehouse, wat de onderhoudbaarheid niet ten goede komt. Ook de nieuw gevonden meetwaarde rond Data Management is (nog) weinig van toepassing bij de deelnemende organisaties. In de interviews is dit beeld bevestigd. Er zijn veel plannen rond de inrichting van meta data, master data en Data Governance, maar in de praktijk heeft dit nog niet geleid tot actie, behalve bij case 4 en 8. Dit is te verklaren, doordat het bij case 4 om een financieel data warehouse gaat, van waaruit verplichte financiële rapporten opgemaakt worden, die rechtstreeks naar externe

toezichthouders gestuurd worden. Bij case 8 is de deelnemer werkzaam bij de afdeling Data Management, wat een verklaring is voor de relatief hoge score.

4.3 Ontwikkelingen

Wat zijn de genoemde ontwikkelingen of alternatieven voor een data warehouse? Is de opkomst van Big Data, Data Science, Internet of Things (IOT), Data virtualisatie of Cloud computing een aanvulling of een bedreiging voor het data warehouse? Figuur 4.16 toont duidelijk de opkomst en populariteit van de Big Data in de Google zoekmachine.

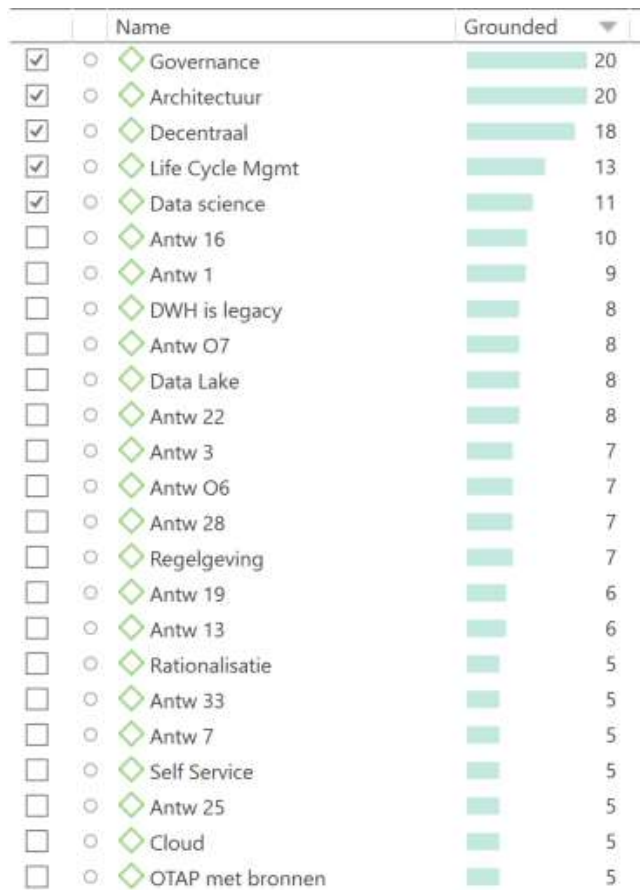


Figuur 4.16 Aantal hits op data warehouse en Big Data in de periode 2004 - 2019. Bron: Google Trends

In de interviews en bij de laatste twee open vragen in de enquête, is ingegaan op de ontwikkelingen in de organisatie die van invloed (kunnen) zijn op het huidige data warehouse landschap. Nadat de transcripties van de interviews in ATLAS.ti zijn ingelezen, samen met de open vragen uit de enquête, zijn de teksten gecodeerd. Bij het coderen werd onderscheid gemaakt in:

1. Bestaande meetwaarden die een (aanvullend) antwoord geven op de vragen uit de enquête, te herkennen aan de codes met "Antw"; en
2. Nieuwe inzichten. Dit zijn onderwerpen (codes) die nu actueel zijn bij Nederlandse financiële dienstverleners en die niet direct gekoppeld kunnen worden aan een bestaande succesfactor of meetwaarde.

In onderstaande frequentieverdeling uit ATLAS.ti worden de belangrijkste codes getoond:



Figuur 4.17 Frequentieverdeling codes uit ATLAS.ti

Figuur 4.17 toont de codes die het meeste genoemd zijn bij de interviews of ingevuld zijn bij de open vragen. De 5 meest genoemde codes representeren de genoemde ontwikkelingen bij de 8 grote financiële dienstverleners in Nederland op dit moment.

Om een beeld te krijgen van de spreiding van deze 5 onderwerpen, is figuur 4.18 uit ATLAS.ti gemaakt. Per interview (case) wordt getoond of (en hoe vaak) het onderwerp besproken werd. Voor de volledigheid zijn ook de enquêteresultaten van de open vragen hieraan toegevoegd:

	DWH enque...	Interview 1	Interview 2	Interview 3	Interview 4	Totals
	2 44	1 60	1 68	1 60	1 45	
Architectuur 20	1		5	7	7	20
Data science 11	2	3	2	1	3	11
Decentraal 18	2	1	10		5	18
Governance 20	3	1	9	4	3	20
Life Cycle... 13	8	1		1	3	13
Totals	16	6	26	13	21	82

Figuur 4.18 Code-Document tabel die de spreiding toont van de 5 onderwerpen per interview. Bron: ATLAS.ti

Te zien is dat Governance en Data science bij alle interviews onderwerp van gesprek was.

De belangrijkste bevindingen worden verder uitgewerkt in de volgende paragrafen.

4.3.1 Governance

Governance, of Data Governance, gaat over het managen van data in een organisatie en maakt onderdeel uit van Data Management (DAMA, 2020). DAMA definieert: "Data Governance –

planning, oversight, and control over management of data and the use of data and data-related resources. While we understand that governance covers ‘processes’, not ‘things’, the common term for Data Management Governance is Data Governance, and so we will use this term”. (Cupoli, 2014, p. 10)

Governance komt als onderwerp in alle interviews ter sprake en ook in de enquêtes wordt het veelvuldig genoemd. Vaak in combinatie met de volgende onderwerpen:



Figuur 4.19 Code Co-Occurance tabel voor Governance. Bron: ATLAS.ti

Governance wordt vaak geassocieerd met de begrippen: Regie, verantwoordelijkheid, eigenaarschap. In figuur 4.19 wordt dit bevestigd. Een data warehouse omgeving kan alleen succesvol zijn, als de directie (antw. 18) en de business (antw. 19) betrokken zijn en verantwoordelijkheid nemen. Een data warehouse is niet alleen een IT-component. De data en de definities zijn niet van IT, net als het beheer van de data niet (alleen) bij IT hoort te liggen (Watson, 2004). In de meeste organisaties wordt dit wel erkend. En een aantal financiële dienstverleners werkt aan een oplossing, maar nergens is dit nog goed geregeld.

De associatie met architectuur en decentraal wordt veroorzaakt doordat er altijd meerdere data warehouses zijn in een organisatie, waardoor een centrale regiefunctie (Data Governance) nog lastig te realiseren is.

4.3.2 Architectuur

Ook architectuur of data-architectuur is onderdeel van DAMA. Hierover wordt gezegd (Cupoli, 2014, p. 10): “Data Architecture – the overall structure of data and data-related resources as an integral part of the enterprise architecture”.

Architectuur is vaak genoemd, maar kent vele contexten:

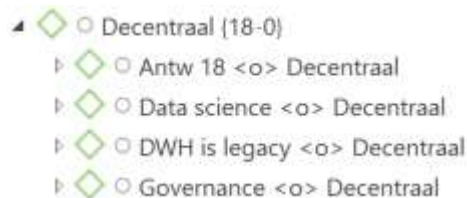


Figuur 4.20 Code Co-Occurance tabel voor Architectuur. Bron: ATLAS.ti

In het algemeen is te stellen dat de onderzochte organisaties nog geen duidelijke referentie-architectuur hebben. Vaak zijn ze er wel mee bezig, maar dan wordt er nog niet op toegezien dat de architectuurprincipes worden nageleefd. Dit heeft raakvlakken met de Data Governance. Ook het decentrale probleem van een versnipperd datalandschap in een organisatie draagt negatief bij aan het naleven van een gemeenschappelijke architectuur.

4.3.3 Decentraal

In alle onderzochte organisaties speelt het probleem van een versnipperd datalandschap. Gemiddeld waren er 8 data warehouses per organisatie (zie paragraaf 4.1), waarbij 3 het minste was. Dit betekent dat een Enterprise Data Warehouse (EDW), zoals beschreven in de jaren 90 (Inmon W. , 1996) met “One version of the truth”, nooit in de praktijk is toegepast bij grote financiële organisaties in Nederland. Ook de gedachte om “Conformed Dimensions” (Kimball, 1996) toe te passen op meerdere data-omgevingen is niet van de grond gekomen.



Figuur 4.21 Code Co-Occurance tabel voor Decentraal. Bron: ATLAS.ti

Wel dient vermeld te worden dat dit probleem breed onderkend wordt en een aantal financiële dienstverleners bezig zijn met het herinrichten van hun datalandschap. Dit ook vanwege de opkomst van Data Science (paragraaf 4.3.5), waarbij het data warehouse dan soms als “Legacy” wordt aangeduid (paragraaf 4.3.4).

4.3.4 Life Cycle Management, Rationalisatie en Legacy

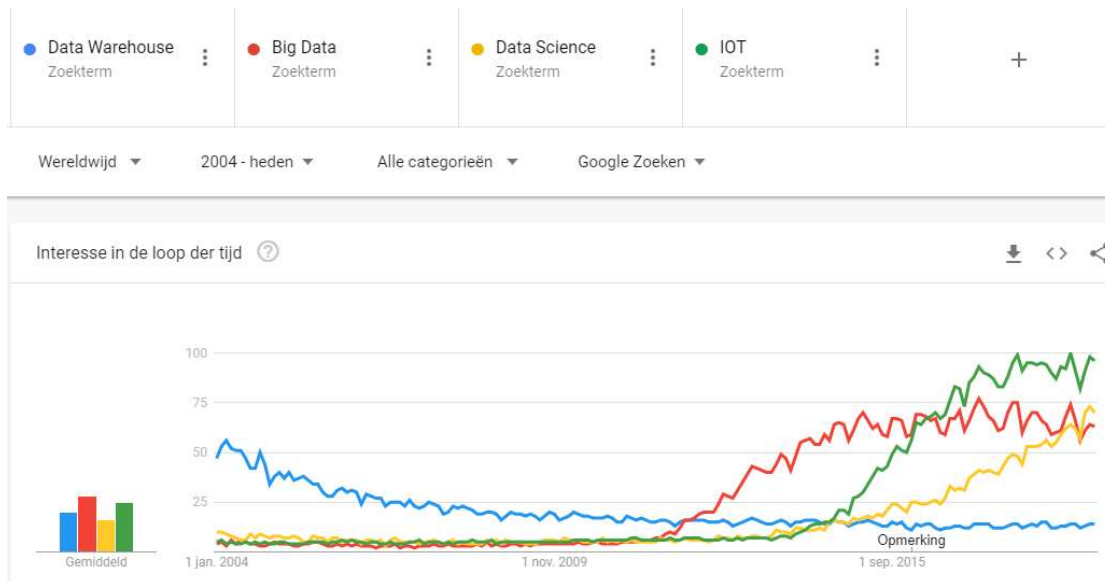
Het begrip Life Cycle Management (LCM) is specifiek bij één case genoemd en ging over het vervangen van een oud data warehouse voor een nieuwe. Bij veel andere cases speelde dit ook, om verschillende redenen. Soms was er een nieuw ERP-systeem op komst en deed het daaraan gelieerde data warehouse mee in de conversie / migratie. Soms ging het over Rationalisatie van het data- of applicatielandschap. En in een paar gevallen werd het data warehouse beschouwd als Legacy en waren de organisaties bezig met een Data Lake of een Data Hub.



Figuur 4.22 Code Co-Occurance tabel voor LCM, Rationalisatie en Legacy. Bron: ATLAS.ti

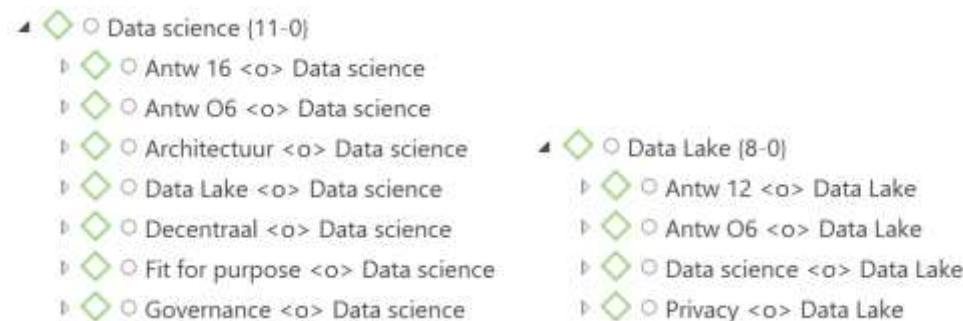
4.3.5 Data Science en Data Lake

Data Science en het bijbehorende Data Lake kan zich verheugen in een enorme opkomst. Figuur 4.23 toont een duidelijke trend in Google de laatste 16 jaar.



Figuur 4.23 Aantal hits per zoekcriterium in de periode 2004 t/m 2019. Bron: Google Trends

Ook bij alle onderzochte organisaties speelde dit onderwerp. En ook alle organisaties hebben naast het data warehouse een Data Lake omgeving ingericht. Vaak worden de data gedeeld, maar soms zijn het twee separate werelden met ieder hun eigen doelstellingen en te ondersteunen business(processen). De intentie om meer en beter samen te werken is er, maar dit heeft nog niet geleid tot concrete actie(s). Er zijn 2 cases bekend waar een nieuwe hybride architectuur (Microsoft, 2016) ingericht wordt, die in de toekomst zowel gestructureerde rapportage vanuit een data warehouse, als Data Science vanuit een Data Lake, moet ondersteunen.



Figuur 4.24 Code Co-Occurance tabel voor DataScience en Data Lake. Bron: ATLAS.ti

4.3.6 Overige ontwikkelingen

Welke overige ontwikkelingen zijn er te melden vanuit de open vragen uit de enquête en de interviews?

Veel onderzochte organisaties zijn bezig te **bezuinigen** of te **reorganiseren**. Rationalisatie en daarmee kostenverlaging van het applicatielandschap speelde hierbij vaak een rol. Het is positief om te horen dat het data warehouse in alle gevallen hierbij betrokken was, ook als het om ERP-applicaties ging. Dat maakt dat data warehouses bij grote financiële dienstverleners voldoende serieus genomen worden.

Cloud kwam regelmatig ter sprake, maar was nooit een factor van discussie. Cloud wordt gezien als infrastructuur en positief geassocieerd met schaalbaarheid en kostenbesparing. Cloud zelf lijkt daarmee geen voorwaarde voor succes.

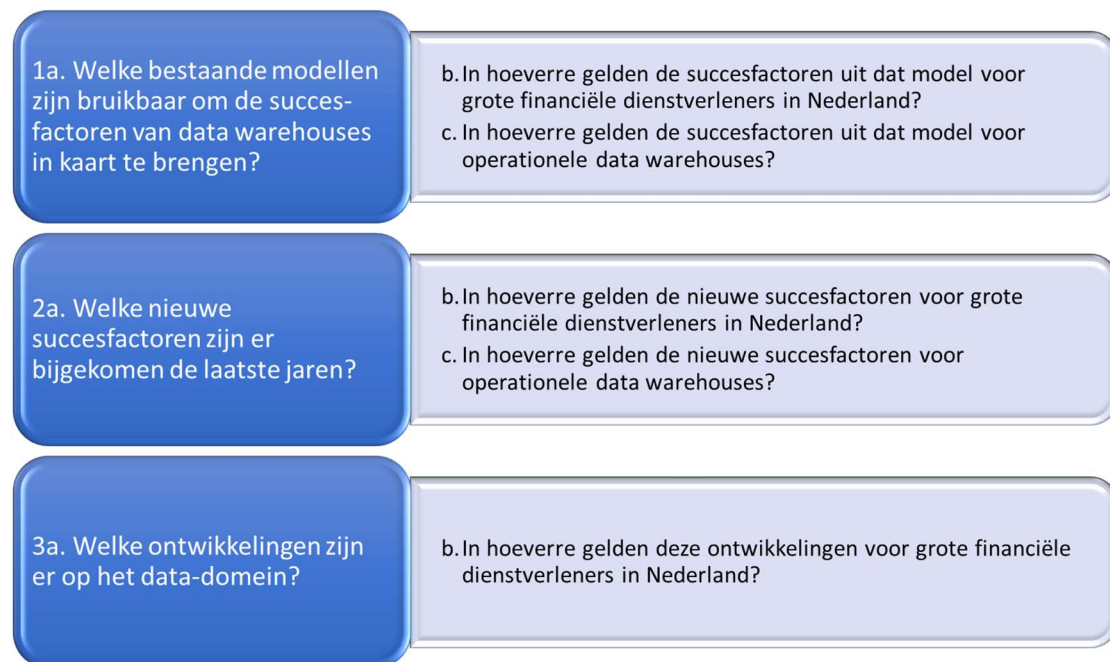
Ook **compliance, security en privacy** kwamen regelmatig ter sprake. Een data warehouse wordt gezien als zeer belangrijk om deze processen te ondersteunen, maar net als alle andere bedrijfsprocessen is dit niet een succesfactor die specifiek benoemd wordt.

Een regelmatig terugkerend onderwerp tijdens de interviews was de kwaliteit en consistentie van de aanleverende **bronsystemen**. Soms ook een persoonlijke frustratie van de geïnterviewden. Bronsystemen doen idealiter aan **release management** en hebben een ontwikkel-, test- en/of acceptatie-omgeving (**OTAP**). Het data warehouse wordt dan niet verrast als bronsystemen wijzigingen doorvoeren, die het data warehouse niet kent, waardoor de load vastloopt. Release management en OTAP is geen voorwaarde voor succes. Het is een middel om de consistentie en betrouwbaarheid van de aanleverend bronsystemen te vergroten, als onderdeel van de succesfactor beheer.

5 Conclusies

5.1 Deelvragen en hoofdvraag

In dit hoofdstuk worden eerst de deelvragen beantwoord vanuit de resultaten van de literatuurstudie, de enquête en de interviews. Er is daarna voldoende informatie aanwezig om de hoofdvraag te beantwoorden: “Wat zijn de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland”.



Figuur 5.1 Deelvragen data warehouse succesfactoren

5.1.1 Deelvragen 1a, 1b en 1c

1a: Tijdens de literatuurstudie zijn er drie bruikbare onderzoeken en modellen gevonden voor data warehouse succes. De modellen zijn gebaseerd op een veelvoud aan interviews, onderzoeken en publicaties, waarbij ze meetwaarden hebben verzameld die de basis waren voor de succesfactoren in het model. De meetwaarden uit de drie modellen zijn vergeleken en geïntegreerd tot één samengesteld model. De meetwaarden zijn geplaatst in een dataverzameltabel die de basis vormde voor de beantwoording van de andere deelvragen. Daarna zijn alle succesfactoren bevestigd door middel van een literatuurstudie van relevante artikelen van de afgelopen 10 jaar.

1b en 1c:

Het samengestelde succesmodel en de meetwaarden in de dataverzameltabel zijn één-op-één vertaald naar vragen in de online enquête over operationele data warehouses. Uit een representatieve steekproef onder 8 grote financiële dienstverleners in Nederland, met een totale populatie van 25 (CBS, 2020), zijn beide deelvragen tegelijk getoetst en kwam het volgende resultaat in figuur 5.2:

Factor / Categorie	Nr	Meetwaarde / Beschrijving	% (N=8)
Waargenomen voordelen	V1	Ondersteunen primaire, secundaire en innovatieve processen	87,50%
	V2	Ondersteunen van strategische doelstelling	87,50%
	V3	Verbeterde individuele productiviteit	62,50%
	V4	Verbeterde besluitvorming op strategisch en tactisch niveau	87,50%
	V5	Verbeterde besluitvorming op operationeel niveau	87,50%
	V6	Ondersteunen van analyse-doeleinden	75,00%
Systeemkwaliteit	S1	Makkelijk in het gebruik en gebruiksvriendelijk	62,50%
	S2	Eenvoudig en efficiënt	50,00%
	S5	Security en autorisatie is goed geregeld	100,00%
	S6	Integratie van meerdere bronsystemen	100,00%
	S7	Goede beschikbaarheid	100,00%
Datakwaliteit	D1	Hoogwaardige datakwaliteit	87,50%
	D2	Voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens	75,00%
	D3	Toereikende actualiteit van data (laadfrequentie en verversingsgraad)	75,00%
	D4	Toegang tot geïntegreerde en consistente data	75,00%
Organisatorische factor	O1	Doelstellingen zijn bekend en voldoen aan de business behoeften	87,50%
	O2	Management vanuit de business en IT is betrokken en ondersteunt	75,00%
	O3	Business (eind)gebruikers zijn actief betrokken	75,00%
	O4	Goede communicatie tussen IT en business	100,00%
	O5	Resources zijn beschikbaar gesteld	75,00%
	O6	Er is een visie en roadmap	87,50%
Technische factor	T3	Goede ontwikkelomgeving	100,00%
	T4	Kundig IT-personeel	100,00%
	T6	Veel data en bronsystemen, waardoor technisch complex	100,00%
Economische factor	E1	Voldoende funding	75,00%
	E2	Voordelen zijn bekend en worden financieel gemeten	50,00%
	E3	Grootte organisatie	>1.500
Beheerfactor	B3	Wijzigingen worden goed uitgevoerd en gemanaged	62,50%
	B4	Goede responsetijd door schaalbaarheid	62,50%
	B5	Gestandaardiseerd, gedocumenteerd en onderhoudbaarheid	25,00%
	B6	Goede datakwaliteit aanleverende bronsystemen	50,00%

Figuur 5.2 Resultaten enquête per meetwaarde in dataverzameltabel

Eén onderdeel scoort onder de 50%, waardoor gesteld kan worden dat deze meetwaarden niet van toepassing is op de operationele data warehouses bij financiële dienstverleners in Nederland. Dit gaat om een meetwaarde bij de succesfactor beheer: standaardisatie en documentatie. Uit de interviews is gebleken dat er niet getwijfeld wordt aan de nut en noodzaak van goede standaardisatie en documentatie, maar dat er in de meeste gevallen geen tijd voor was (gemaakt).

Ook vallen er drie onderdelen op met een 50% score. Een data warehouse staat vanuit de business niet bekend als een eenvoudig en efficiënt systeem. Goede BI-tools kunnen het leed verzachten, maar dat neemt niet weg dat een data warehouse in de kern een complex systeem is, wat in alle onderzochte gevallen door IT wordt beheerd.

Die complexiteit wordt ook bevestigd door de matige score van datakwaliteit van de bronsystemen. Uit de interviews is dit probleem vaak naar voren gekomen. Niet alleen de kwaliteit van data, maar ook de consistentie van aanlevering en release management met een goed werkende OTAP-omgeving, is vaak de oorzaak van problemen.

Een ander punt van zorg is de 50% score bij het financieel meten van de voordelen bij data warehouses en/of data warehouse-(verbeter)projecten. Als deze voordelen onvoldoende financieel worden onderbouwd, kan op termijn, bijvoorbeeld bij rationalisatie, het bestaansrecht van een data warehouse in twijfel worden getrokken.

5.1.2 Deelvragen 2a, 2b en 2c

2a: Deelvraag 2a behandelt nieuwe gevonden succesfactoren in de literatuur van de laatste 10 jaar. Er zijn een drietal onderwerpen aan het licht gekomen in combinatie met data warehousing:

1. Een **Agile werkwijze** om het bestaande data warehouse stap-voor-stap te verbeteren;
2. **Data Management**, een vakgebied (en verzamelbegrip) in opkomst en waar het data warehouse onderdeel van uitmaakt;
3. Een nieuwe **succesfactor: Beheer**, die ook de vorige twee begrippen omvat. Beheer als succesfactor ontbrak tot nu toe, omdat vrijwel alle bestaande modellen uitgaan van data warehouse-projecten en -implementaties.

2b en 2c: Net als de bestaande meetwaarden zijn ook de drie nieuwe onderwerpen vertaald in vragen in de enquête en besproken in de interviews met grote financiële dienstverleners in Nederland. Omdat de drie nieuwe meetwaarden horen bij de nieuwe succesfactor: beheer, worden de deelvragen 2b (financiële dienstverleners) en 2c (operationele data warehouses) tegelijk beantwoord. Het resultaat:

Factor / Categorie	Nr	Meetwaarde / Beschrijving	% (N=8)
Beheerfactor	B1	Verbeterpunten in de vorm van een backlog zijn bekend en geregistreerd	75,00%
	B2	Continue verbeteringen door Agile werkwijze	75,00%
	B7	Data Management is ingericht	37,50%

Figuur 5.3 Resultaten enquête beheerfactor in dataverzameltabel

De **Agile werkwijze** is ondervraagd in B1 en B2 en in de interviews ter sprake gekomen. 75% van de respondenten gaven aan Agile te werken (meestal Scrum) en ook in de interviews kwam dit beeld naar voren. Het is belangrijk om Agile development vanuit IT in te zetten, in combinatie met “klassiek” projectmanagement, dat vanuit de business en het management geïnitieerd en gefund worden (Batra, 2017). De combinatie geeft het meeste kans op succes.

Data Management is uitgevraagd in B7 en in de interviews veelvuldig ter sprake gekomen. Ondanks een lage score van 37,5% op de onderdelen metadata en masterdata, wordt er wel veel gesproken over Data Governance bij de grote financiële dienstverleners in Nederland. Hierover meer in paragraaf 5.1.3.

Als succesfactor is het onderwerp “**Beheer**” minder sterk van toepassing op de doelgroep, dan de andere factoren. Dit betekent niet dat de factor minder belangrijk wordt geacht. Het beheer van een data warehouse is een continu proces; het blijft groeien in datavolume, wat kan leiden tot problemen met beheer en onderhoud (Huang, 2007). Ook tijdens de interviews kwam dit beeld naar voren. Beheer is belangrijk, maar bij een aantal respondenten wordt er vol ingezet op nieuwe alternatieve oplossingen, waardoor het oude data warehouse “end-of-life” is, wat direct ten koste van het beheer gaat.

5.1.3 Deelvragen 3a en 3b

Als er gekeken wordt naar de ontwikkelingen in de markt op het gebied van data en naar de activiteiten op dat vlak bij de grote financiële dienstverleners, dan staat Data Management wederom, of nog steeds, hoog op de agenda. Als gekeken wordt naar het hele spectrum van DMBOK (DAMA, 2020), dan is er één onderwerp dat het meest genoemd wordt tijdens de enquête en interviews: Data Governance.

Dit onderwerp wordt geassocieerd met:

1. **Decentraal.** Grote organisaties hebben meerdere data warehouses voor meerdere business processen. Vaak zijn Finance en Marketing enthousiaste afnemers van data uit een data warehouse, maar van samenwerking is nauwelijks sprake. De verantwoordelijkheid en eigenaarschap van de data is dus ook versnipperd in de organisatie en meestal lost IT dit probleem “tijdelijk” op. Het idee dat een Enterprise Data Warehouse (EDW) de “Single version of the truth” toont (Inmon B. , 2004) gaat dus nooit op in dit soort grote organisaties;
2. **Architectuur.** Alle geïnterviewden gaven aan dat hun organisatie bezig was met het opzetten van een nieuwe data warehouse-, Data Lake- of hybride architectuur. Dit initiatief gaat altijd gepaard met de intentie tot het verbeteren of inrichten van Data Governance. Zaken als nauwere samenwerking tussen IT en Business, tussen BI en Data Science teams, tussen bronsystemen en rapportenbouwers. Alle vormen zijn langsgelopen in de gesprekken. De intenties zijn goed, maar concreet is er nog niks gerealiseerd;
3. **Data Science.** Alle onderzochte organisaties hebben naast het data warehouse een Data Lake omgeving ingericht. Vaak wordt wel de data gedeeld, maar soms zijn het twee separate werelden met ieder hun eigen doelstellingen en te ondersteunen business(processen). De intentie om meer en beter samen te werken is er, maar dit heeft nog niet geleid tot concrete actie(s). Qua Data Governance nog een hele uitdaging.

Bovenstaande onderwerpen staan op de agenda van de onderzochte organisaties en versterken de centrale positie van een data warehouse als bron voor structurele rapportage en analyse. Ondanks de relatieve ouderdom en complexiteit, is er nog lang geen sprake van een “end-of-life” situatie en blijft investeren in een modern data warehouse noodzakelijk.

5.2 Onderzoeksvraag

Het antwoord op de onderzoeksvraag:

Wat zijn de succesfactoren van operationele data warehouses bij grote financiële dienstverleners in Nederland?

De directe succesfactoren zijn:

1. Waargenomen voordelen, zoals het ondersteunen van processen en besluitvorming om daarmee strategische doelen te bereiken;
2. Systeemkwaliteit, zoals gebruiksvriendelijk, efficiënt en beschikbaar;
3. Datakwaliteit en informatieve waarde.

De indirecte factoren die dit mogelijk moeten maken, zijn:

4. Organisatorische factoren, zoals business betrokkenheid en management support;
5. Technische factoren, zoals een goede ontwikkel-tools en kundig IT-personeel;
6. Economische factoren, zoals voldoende budget;
7. Beheerfactoren, zoals continu verbeteren, schaalbaarheid en Data Management.

Waar bovenstaande factoren nog redelijk overeenkomen met de succesfactoren van “gewone” Information Systems (IS), is er ook duidelijk een onderscheid:

1. Datakwaliteit is bij een data warehouse cruciaal. Een business gebruiker moet 100% kunnen vertrouwen op de juistheid, volledigheid en actualiteit van de data waarop hij rapporteert en beslissingen neemt.

2. Beheerfactoren zijn bij een data warehouse anders dan andere IT-systemen. De data van bronsystemen moet consistent zijn en hoe beter Data Management voor de hele organisatie ingericht is, hoe meer een data warehouse daarvan profiteert.

Bij financiële dienstverleners in Nederland lijkt de datakwaliteit van voldoende hoog niveau (87,5%), maar zijn de beheerfactoren nog niet voldoende op orde, met een lage score voor standaardisatie, documentatie en Data Management.

5.3 Probleemomschrijving

Als gekeken wordt naar de probleemstelling zoals beschreven in paragraaf 1.2, dan is het volgende te concluderen:

Complexiteit & Kosten: Het onderhouden en uitbreiden van een data warehouse is een continu proces. Waar begin deze eeuw de effort in tijd en geld ging zitten in lange data warehouse implementatie-projecten, is er nu de focus op beheer en onderhoud. Met een Agile-werkwijze zijn de complexiteit en de kosten beter in de hand te houden. Zodra Enterprise Data Management in al zijn aspecten uitgerold wordt bij de grote financiële dienstverleners, zal het beheer en onderhoud van een data warehouse veel minder complex worden. De onderzochte organisaties onderkennen dit, maar doen er nog weinig aan.

Vertrouwen en adoptie vanuit de business: Een data warehouse is inmiddels niet meer weg te denken bij grote financiële dienstverleners. De nut en noodzaak is bewezen in de vorm van geïntegreerde financiële dashboards, klant-analyses en verplichte rapportages naar externe toezichthouders (compliance). Door het succes ervan, is er een nieuw probleem bijgekomen: Elke business heeft zijn eigen data warehouse en de samenwerking daartussen schiet ernstig tekort. Een goed ingerichte Data Governance, als onderdeel van Data Management, wordt onderkend als oplossing, maar is nog niet geïmplementeerd bij de onderzochte organisaties. Hier is de uitdaging om het topmanagement te overtuigen, onder andere door Business en IT samen op te laten trekken.

Alternatieve oplossingen: Data Science als nieuw analytisch business proces, is bij alle onderzochte organisaties gaande. De data die hiervoor nodig is komt uit een Data Lake, die in enkele gevallen samenwerkt met het data warehouse. Ook hier wordt integratie door een gedegen data-architectuur onderkend, maar in de praktijk blijken de onderzochte organisaties (op een enkele na) hier nog geen voortgang te boeken.

6 Discussie en aanbevelingen

In dit hoofdstuk wordt gekeken naar de afbakening van het onderzoek en de beperkingen die tijdens het onderzoek geconstateerd zijn. In paragraaf 6.2 wordt een voorzet gegeven voor vervolgonderzoek op basis van de onderzoeksresultaten en de beperkingen en in paragraaf 6.3 volgen de aanbevelingen voor de doelgroep in de praktijk.

6.1 Afbakening en beperkingen

Literatuurstudie

De bestaande succesfactoren zijn tot stand zijn gekomen in de jaren nul van deze eeuw met de modellen van Wixom et al. (2001), Hwang et al. (2008) en Ramamurty et al. (2007). Om de nieuwe succesfactoren te ontdekken is daarom gekozen voor de periode 2009 tot heden. De bestaande modellen kennen een aantal beperkingen:

1. Veruit de meeste publicaties komen uit de Verenigde Staten, waar ook de theorie in de jaren 90 ontstaan is. Ook de gebruikte succesmodellen in deze thesis komen uit de Verenigde Staten;
2. In de literatuurreferenties van de bestaande modellen in de jaren nul van deze eeuw wordt de nadruk gelegd op data warehouse succes bij implementaties. Het operationele karakter van een data warehouse (onder andere beheer en onderhoud) speelt nagenoeg geen rol in de totstandkoming van het basismodel. Dit wordt onderkend in dit onderzoek;
3. Een aantal bronnen die gehanteerd zijn bij de bestaande modellen gingen over IT-succes in het algemeen en niet over data warehouse-succes in het bijzonder;
4. In de literatuur van de afgelopen 10 jaar is gezocht naar de begrippen "Data" AND ("Warehouse" OR "Warehousing") in de titel. Als gezocht wordt op deze woorden binnen het artikel, kwamen er meer dan 5.000 hits terug die (steekproefsgewijs) niets te maken hadden met het onderzoek.

Ondanks en dankzij deze beperkingen biedt dit onderzoek en het samengestelde succesmodel een prima basis voor actualisatie in deze thesis. Met name de laatste 10 jaar is er veel meer bekend geworden over data warehouse successen of tekortkomingen.

Model

Bij het analyseren en integreren van de 3 modellen zijn de begrippen: Schaalbaarheid en Responstijd samengevoegd tot één meetwaarde. Dit is achteraf niet verstandig geweest, omdat schaalbaarheid iets zegt over het (flexibel) beheren van een data warehouse, terwijl responstijd meer over de systeemkwaliteit gaat. In de enquête is dit wel apart ondervraagd: vraag 2 gaat over de efficiëntie en snelheid van het systeem en vraag 30 gaat over schaalbaarheid.

Enquête en interviews

Een enquête of interview kan vooroordelen of vooringenomenheid van de onderzoeker bevestigen, zoals staat beschreven in het boek over Case Study Surveys (Chmiliar, 2012). De vragen uit de enquête zijn gebaseerd op het samengestelde succesmodel en kunnen in die zin dus geen uitslag beïnvloeden. Wel heeft de onderzoeker tijdens de semigestructureerde interviews het onderwerp van gesprek kunnen beïnvloeden. De geïnterviewden hebben zelf kunnen aangeven of ze alle relevante informatie in hun ogen hebben doorgegeven. Ook de enquête gaf ruimte voor opmerkingen en persoonlijke inbreng.

Functie

De functie van de deelnemers aan de enquête of het interview, is van invloed gebleken op het resultaat. Bij het beantwoorden van de vragen zijn de IT-ers en functioneel beheerders van een data warehouse, de mensen die dus dicht op de materie zitten, overwegend positief over de succesfactoren van “hun” data warehouse. De waargenomen voordelen staan verder van ze af, maar ook die schatten ze vrij positief in. Als men kijkt naar de business functies en architectuur, dan zijn de geluiden meer kritisch. Vaak wordt er uitgekeken naar een TO-BE-architectuur met moderne mogelijkheden en wordt het “oude” data warehouse wat negatiever afgeschilderd. Bij een aantal cases hebben meerdere functies de enquête ingevuld. Hun antwoorden zijn in het resultaat per case soms genivelleerd tot een neutrale “Van toepassing”.

Deelvragen

Bij het verwerken van de resultaten en conclusies is gebleken dat alle onderzochte organisaties binnen de doelgroep meerdere operationele data warehouses hadden. Daarmee kon antwoord gegeven worden op de deelvragen 1b en 1c en op 2b en 2c. Achteraf bezien had het opsplitsen van de b-vragen (grote financiële dienstverleners) en de c-vragen (operationele data warehouses) geen toegevoegde waarde.

6.2 Vervolgonderzoek

Alle onderzochte organisaties hadden meerdere data warehouses in beheer. De komst van Data Science en een nieuwe hybride architectuur, zal deze decentrale situatie niet snel wegnemen, eerder versterken. Organisaties hebben hier dus mee te leren leven. Data Management in het algemeen en Data Governance in het bijzonder, kan hen hierbij helpen. Vervolgstudie is noodzakelijk om te onderzoeken in welke mate Data Management hier een bijdrage kan leveren en welke succesfactoren en meetwaarden hierbij geraakt worden.

Naast grote financiële dienstverleners met meerdere data warehouses, is het interessant om andere typen bedrijven te onderzoeken. De middelgrote ondernemingen in Nederland, die vaak later begonnen zijn aan data integratie, hebben meestal te maken met één centraal data warehouse met één centrale beheerorganisatie. Dit kan effect hebben op de succesfactoren en de theorie zoals die in de jaren 90 opgesteld is.

6.3 Aanbevelingen voor de praktijk

Het doel van dit onderzoek was om succesfactoren te benoemen voor bestaande operationele data warehouses bij financiële dienstverleners in Nederland. Vanuit de theorie en de literatuur is een bruikbaar succesmodel opgesteld, waar elke organisatie mee aan de gang kan.

Uit dit onderzoek is gebleken dat het beheer van een data warehouse bij deze organisaties nog vaak te wensen overlaat. Het advies is dan ook: Investeer in het data warehouse door het verbeteren van de standaardisatie en de documentatie. Dit heeft een positief effect op de onderhoudbaarheid van het data warehouse en, op termijn, betere samenwerking met andere data warehouses binnen de organisatie met een mogelijk synergievoordeel op beheer.

Een ander punt van zorg is de consistente datakwaliteit van de aanleverende bronsystemen. Data Management of beter: Enterprise Data Management, kan hier veel betekenen. Data Management met al zijn onderdelen is geen luxe of een nice-to-have, het is pure noodzaak voor grote organisaties met een versnipperd datalandschap en onvoldoende Data Governance.

Onderdeel van Data Management is ook een gedegen data-architectuur. In alle onderzochte organisaties is gestart met Data Science, vaak met een achterliggende Data Lake. Deze Data

Science teams ondersteunen andere business processen, dan de standaard rapportages uit “klassieke” data warehouses. Maar het betekent feitelijk weer een “data-eilandje” erbij. Een toekomstvaste data-architectuur met principes en modellen, is noodzakelijk om data-integratie binnen de organisatie te bevorderen. Daarmee geeft het ook direct een aanleiding om vanuit de business meer en beter te gaan samenwerken op het gebied van informatiemanagement.

Om alle decentrale oplossingen zoveel mogelijk samen te brengen, verdient het dus de aanbeveling om zo snel mogelijk te beginnen met (Enterprise) Data Management. Door te investeren in: Datakwaliteit, data-architectuur, metadata, Reference- en Masterdata, data modelering én Data Governance, krijgt men een positief effect op alle succesfactoren van een data warehouse, maar in het bijzonder de: Beheerfactoren.

7 Referenties

Hieronder staan de referenties van de boeken, artikelen en websites die in deze thesis gebruikt zijn.

- Al-Debei, M. (2011). Data warehouse as a backbone for business intelligence: Issues and challenges. *European Journal of Economics, Finance and Administrative Sciences*, 33, 153-166.
- Alhyasat, B. (2013). Data Warehouse Success and Strategic Oriented Business Intelligence: A Theoretical Framework. *Journal of Management Research*.
- Almabhouh, A. S. (2011). Examining the Influence of Relationship Quality on Data Warehouse Success. *International Journal of Modeling and Optimization*.
- Alnassar. (2014). Challenges in the Successful Implementation of Data Warehouse. *Journal of Management Research*.
- Ariyachandra, T. W. (2010). Key organizational factors in data warehouse architecture selection. *Decision Support Systems* 49.
- ATLAS.ti. (2020). Opgehaald van ATLAS.ti: <https://atlasti.com/>
- Batra, D. (2017). Adapting agile practices for data warehousing. *Journal of Database Management (JDM)* , v28 n4 (10 2017): 1-23.
- Bloomberg, L. V. (2008). *Completing Your Qualitative Dissertation: A Roadmap From Beginning to End*. Los Angeles: SAGE publications.
- Bryman, A. &. (2015). *Business Research Methods*. . United Kingdom: Oxford.
- Burns, A. B. (2008). *Basic Marketing Research*. New Jersey: Pearson Education (page 250).
- Burtescu, C. &. (2010). Data warehouse and data mining - databases next step. Annals of the University of Oradea. *Economic Science*, 1, 804-809.
- CBS. (2020). CBS StatLine. Opgehaald van opendata.cbs.nl: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81589ned/table?dl=2FAD8&ts=1581774536691>
- Chaudhuri, S. D. (2011). An Overview of Business Intelligence Technology. *Communications of the ACM*, 88-98.
- Chmiliar, L. (2012). Case Study Surveys. *Encyclopedia of Case Study Research - Sage researchmethods*.
- Cooper, B. W. (2000). Data Warehousing Supports Corporate Strategy at First American Corporation. *MIS Quarterly*.
- Cupoli, P. (2014). *DAMA-DMBOK2 Framework*. USA: DAMA.
- Curtis, K. J. (1999). Issues in Building a Successful Data Warehouse. *Information Strategy*.
- DAMA. (2020). Opgehaald van DAMA International: <https://dama.org/content/body-knowledge>
- Difference, I. (2015, Jan). *Is the data warehouse dead?* Opgehaald van TDWI.org: tdwi.org/media
- Easion.nl. (2020). Opgehaald van Easion Survey: <https://apps.parantion.nl/easionx/#/login>
- Eisenhardt, K. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4), 532-550.
- Foley, J. (1997). Data warehouse pitfalls. *Informationweek*, 631.
- Gartner. (2018). *Logical Data Warehouse*. Opgehaald van www.gartner.com: <https://www.gartner.com/imagesrv/media-products/pdf/solix/Solix-1-4WFTJQG.pdf>
- Goldstein, H. (2005). Who Killed the Virtual Case File? <https://spectrum.ieee.org/computing/software/who-killed-the-virtual-case-file>.
- Goodwin, B. (2011, Jan). *Computerweekly.com*. Opgehaald van Computerweekly.com: <https://www.computerweekly.com/news/1280094776/Poor-communication-to-blame-for-business-intelligence-failure-says-Gartner>
- Google. (2020). Opgehaald van Google Scholar: <https://scholar.google.nl/>
- Henschen, D. (2008). The data warehouse revised. *Informationweek*, 1188.

- Huang, S. C. (2007). Data warehouse enhancement: A semantic cube model approach. *Information Sciences*, 177(11), 2238-2254.
- Hwang, M. &. (2008). A structural model of data warehousing success. *Journal of Computer Information Systems*, 49(1), 48-56.
- Inmon, B. (2004). The Single Version Of The Truth. *BI Network*.
- Inmon, W. (1996). The data warehouse and data mining. *Communications of the Acm*, 39 (11).
- Inmon, W. (2005). *Building the Data Warehouse, Fourth Edition*. John Wiley & Sons.
- Isik, O. &. (2013). Business intelligence success: The roles of BI capabilities and decision. *Information & Management (Elsevier)*, 13-23.
- Jamieson, S. (2004). A simple and reliable method of scoring the thurstone attitude scales. *Blackwell Publishing*.
- Kimball, R. (1996). *The Data Warehouse Toolkit*. John Wiley & Sons.
- Kimball, R. R. (2016). *The Kimball Group Reader*. New York: Wiley.
- Krishnamurthi, M. (2014). Increasing data warehousing success rates a tale of two companies. *. EJM, Volume 14, nr 1*.
- Lans, R. v. (2011). Niemand wil een datawarehouse. *Database Magazine*, Nr 1.
- Linstedt, D. (2015). *Building a Scalable Data Warehouse with Data Vault 2.0*. Morgan Kaufmann.
- Lokke, A. S. (2014). Theory Testing using Case Studies. *The Electronic Journal of Business Research Methods*.
- March, S. &. (2007). Integrated decision support systems: A data warehousing perspective. *Decision Support Systems*, 43(3), 1031-1043.
- Microsoft. (2016). *Microsoft Modern Data Warehouse white paper*. Opgehaald van Microsoft.com: http://download.microsoft.com/download/C/2/D/C2D2D5FA-768A-49AD-8957-1A434C6C8126/Microsoft_Modern_Data_Warehouse_white_paper.pdf
- Morris, J. K. (2013). CRITICAL SUCCESS FACTORS FOR DATA WAREHOUSING: A CLASSICAL ANSWER TO A MODERN QUESTION. *Issues in Information Systems*, Volume 14, Issue 1, pp.376-384.
- Mosley, M. B. (2010). *The DAMA Guide to the Data Management Body of Knowledge*. Bradley Beach, New Jersey: Technics Publications.
- Nelson, J. (2013, April 22). Opgehaald van CX Unplugged: https://cxounplugged.com/2013/04/shadow_it/
- Oktavia, T. (2014). Implementing data warehouse as a foundation for decision support system. *Journal of Theoretical and Applied Information Technology*.
- Oost, H. &. (2016). *Een onderzoek voorbereiden*. Amersfoort: ThiemeMeulenhof.
- Paré, G. (2004). Investigating Information Systems with Positivist. *Communications of the Association for Information Systems*.
- Perez, J. B. (2009). A relevance model for a data warehouse contextualized with documents. *Information Processing and Management*.
- Raden, N. (2019). The Data Warehouse is dead - long live the Data Warehouse. *Diginomica*.
- Ramamurthy, K. &. (2007). An empirical investigation of the key determinants of data warehouse adoption. *ScienceDirect / Elsevier Decision Support Systems*.
- Rockart, J. &. (1981). *A primer on critical success factors*. Cambridge: MIT.
- Schiff, M. (2019). Long Live the Traditional Data Warehouse. *TDWI - The Data Warehouse Institute*.
- Sen, A. R. (2012). A model of data warehousing process maturity. *IEEE Transactions on Software Engineering*.
- Serra, J. (2017). *James Serra's Blog*. Opgehaald van Big Data and Data Warehousing: <https://www.jamesserra.com/archive/2017/12/is-the-traditional-data-warehouse-dead/>
- Steekproefcalculator. (2020). *Sample Size Calculator*. Opgehaald van SurveyMonkey: <https://nl.surveymonkey.com/mp/sample-size-calculator/>
- Swan, G. (2011, Feb). *CIO.com*. Opgehaald van CIO.com: <https://www.cio.com/article/2411239/nine-data-warehousing-trends-for-the-cio.html>

- Takecian, P. (2014). Methodological Guidelines and Adaptive Statistical Data. *Simpósio Brasileiro de Bancos de Dados - SBB D 2012*.
- Thorne, S. (2000). The data warehouse: Keeping it simple. mit shares valuable lessons learned from a successful data warehouse implementation. *Educause Quarterly*, 23(3), 26-30.
- Turban, E. (2011). *Business Intelligence: a managerial approach*. Boston: Prentice Hall.
- Watson, H. F. (2004). Data warehouse governance: Best practices at blue cross and blue shield of north carolina. *Decision Support Systems*, 38(3), 435-450.
- Whiting, R. (2003). The data-warehouse advantage. *Informationweek*, 949.
- Winter, R. M. (2001). Organization of data warehousing in large service organizations: A matrix approach based on data ownership and competence centers. *Journal of data warehousing*.
- Wixom, B. &. (2001). An empirical investigation of the factors affecting data warehousing success. *Mis Quarterly*, 25(1).
- Yeoh, W. K. (2008). Managing the Implementation of Business Intelligence Systems. *International Journal of Enterprise Information Systems*.
- Yin, R. K. (2013). *Case Study Research (5th edition)*. California, USA: SAGE Publications, Inc;.

8 Bijlage

8.1 Criteria en zoektermen literatuuronderzoek

Als belangrijkste informatiebron is HUGO gebruikt: Het online bibliotheekstelsel van de Hogeschool Utrecht (HU) die toegang geeft tot wetenschappelijke artikelen, proefschriften en boeken. Ook Google Scholar (Google, 2020) is geraadpleegd, vooral als de gevonden artikelen niet beschikbaar waren via de universitaire kanalen.

De gebruikte parameters bij het zoeken zijn als volgt samengesteld:

Parameter	Waarde
Peer-reviewed	Ja
Bibliotheken	Wereldwijd
Taal	Engels en Nederlands
Jaar	Vanaf 2009
Materiaalsoorten	Alle
Databases	Alle
Auteurs	Alle

Tabel 8.1 Ingestelde parameters voor literatuuronderzoek data warehouse succesfactoren

De gebruikte zoektermen zijn direct afgeleid van de onderzoeksvraag en zijn als volgt:

Zoektermen	# Hits
data AND (warehouse OR warehousing) in tekst	5.021
data AND (warehouse OR warehousing) in titel	699
data AND (warehouse OR warehousing) AND best AND practice in titel	0
data AND (warehouse OR warehousing) AND success in titel	9
data AND (warehouse OR warehousing) AND maturity in titel	1
data AND (warehouse OR warehousing) AND factor in titel	9
data AND (warehouse OR warehousing) AND principle in titel	0
data AND (warehouse OR warehousing) AND starting AND points in titel	0
data AND (warehouse OR warehousing) AND basic in titel	0
data AND (warehouse OR warehousing) AND relevance in titel	1
data AND (warehouse OR warehousing) AND assessment in titel	10
data AND (warehouse OR warehousing) AND failure in titel	0
data AND (warehouse OR warehousing) AND goal in titel	0
data AND (warehouse OR warehousing) AND function in titel	3

Tabel 8.2 Zoekcriteria voor literatuuronderzoek data warehouse succesfactoren

Alternatieve namen voor een data warehouse, zoals: data warehouse, DW, EDW, Information Factory, leverde geen bruikbare output op.
Nederlandse publicaties zijn niet gevonden.

8.2 Publicaties op basis van zoekcriteria

Op basis van de zoekcriteria én indien aanwezig, zijn de volgende publicaties geselecteerd.

Zoekterm: “data warehouse Success”

9 titels gevonden, waarvan een aantal dubbelen en 5 relevant

Nr.	Auteur(s)	Titel	Publicatie
1	Batra, D. (Batra, 2017)	Agile values or plan-driven aspects: Which factor contributes more toward the success of data warehousing, business intelligence, and analytics project development?	The Journal of Systems & Software, 146, 249-262. doi:10.1016/j.jss.2018.09.081
2	Krishnamurthi, M. (Krishnamurthi, 2014)	Increasing data warehousing success rates a tale of two companies.	European Journal of Management, 14(1), 7-12. doi:10.18374/EJM-14-1.1
3	Alhyasat, E., & Al-Dalahmeh, M. (Alhyasat, 2013)	Data warehouse success and strategic oriented business intelligence: A theoretical framework	Journal of Management Research, 5(3). doi:10.5296/jmr.v5i3.3703
4	Alnassar, B. (Alnassar, 2014)	Challenges in the successful implementation of data warehouse.	Journal of Management Research, 6(3), 214-214. doi:10.5296/jmr.v6i3.5196
28	Alaaeddin Almabhouh, Abdul R. Saleh, and Azizah Ahmad (Almabhouh, 2011)	Examining the Influence of Relationship Quality on Data Warehouse Success	International Journal of Modeling and Optimization, Dec 2011

Zoekterm: “data warehouse Maturity”

Nr.	Auteur(s)	Titel	Publicatie
5	Sen, A., Ramamurthy, K., & Sinha, A. (Sen A. R., 2012)	A model of data warehousing process maturity	IEEE Transactions on Software Engineering, 38(2), 336-353

Zoekterm: “data warehouse Factor”

9 titels gevonden, waarvan een aantal dubbelen en 2 relevant

Nr.	Auteur(s)	Titel	Publicatie
6	Ariyachandra, T., & Watson, H. (Ariyachandra, 2010)	Key organizational factors in data warehouse architecture selection	Decision Support Systems, 49(2), 200-212. doi:10.1016/j.dss.2010.02.006
7	Oktavia, T. (Oktavia, 2014)	Implementing data warehouse as a foundation for decision support system (perspective: Technical and nontechnical factors).	Journal of Theoretical and Applied Information Technology, 60(3), 476-482.

Zoekterm: "data warehouse Relevance"

Nr.	Auteur(s)	Titel	Publicatie
10	Pérez, J., Berlanga, R., & Aramburu, M. (Perez, 2009)	A relevance model for a data warehouse contextualized with documents	Information Processing and Management, 45(3), 356-367. doi:10.1016/j.ipm.2009.11.001

Zoekterm: "data warehouse Assessment"

10 titels gevonden, waarvan veel dubbelen en 1 relevant

Nr.	Auteur(s)	Titel	Publicatie
11	Gosain, A., Sabharwal, S., & Nagpal, S. (2011).	Assessment of quality of data warehouse multidimensional model.	International Journal of Information Quality, 2(4), 344-358. doi:10.1504/IJIQ.2011.043782

Overige artikelen (oud en recent) op basis van referenties of eerdere publicaties van relevante onderzoekers

Nr.	Auteur(s)	Titel	Publicatie
12	Ramamurthy, K., Sen, A., & Sinha, A. (Ramamurthy, 2007)	An empirical investigation of the key determinants of data warehouse adoption	Decision Support Systems, 44(4), 817-841. doi:10.1016/j.dss.2007.10.006 - 2008
13	March, S., & Hevner, A. (March, 2007)	Integrated decision support systems: A data warehousing perspective.	Decision Support Systems, 43(3), 1031-1043. doi:10.1016/j.dss.2005.05.029 - 2007
14	Hwang, M., & Xu, H. (Hwang, 2008)	A structural model of data warehousing success.	Journal of Computer Information Systems, 49(1), 48-56. doi:10.1080/08874417.2008.11645305 - 2015
15	W. Inmon (Inmon W. , 1996)	The Data Warehouse and Data Mining	Communications of the Acm, 39 - 1996
16	S. Thorne (Thorne, 2000)	The data warehouse: Keeping it simple. MIT shares valuable lessons learned from a successful data warehouse implementation	Educause Quarterly, 23(3), 26-30 - 2000
17	Whiting, R. (Whiting, 2003)	The data-warehouse advantage	Informationweek, 949(949) - 2003
18	Henschen, D. (Henschen, 2008)	The data warehouse revised	Informationweek, 1188(1188) - 2008
19	Foley, J. (Foley, 1997)	Data warehouse pitfalls	Informationweek, 631(631) - 1997

20	Watson, H., Fuller, C., & Ariyachandra, T. (Watson, 2004)	Data warehouse governance: Best practices at blue cross and blue shield of north carolina.	Decision Support Systems, 38(3), 435-450. doi:10.1016/j.dss.2003.06.001 - 2004
21	Al-Debei, M. (Al-Debei, 2011)	Data warehouse as a backbone for business intelligence: Issues and challenges.	European Journal of Economics, Finance and Administrative Sciences, 33(33), 153-166 – 2011
22	Burtescu, C., & Burtescu, E. (Burtescu, 2010)	Data warehouse and data mining - databases next step.	Annals of the University of Oradea: Economic Science, 1(1), 804-809(2010)
23	Huang, S., Chou, T., & Seng, J. (Huang, 2007)	Data warehouse enhancement: A semantic cube model approach.	Information Sciences, 177(11), 2238-2254. doi: 10.1016/j.ins.2006.12.022 - 2007
24	Wixom. B.H. and Watson. H.J. (2001).	An empirical investigation of the factors affecting data warehousing success.	MIS Quarterly, 25. 1, 17-41.
25	Winter. R., and Meyer. M. (Winter, 2001)	Organization of data warehousing in large service organizations: A matrix approach based on data ownership and competence centers	Journal of Data Warehousing, 6, 4, 23-29.
26	March, S., & Hevner, A. (March, 2007)	Integrated decision support systems: A data warehousing perspective.	Decision Support Systems, 43(3), 1031-1043. doi:10.1016/j.dss.2005.05.029 - 2007
27	D. Batra (Batra, 2017)	Adapting agile practices for data warehousing (eerste 2 pagina's)	Journal of Database Management (JDM) v28 n4 (10 2017): 1-23

Niet wetenschappelijke artikelen

Nr.	Auteur(s)	Titel	Publicatie
29	Gartner (Gartner, 2018)	Logical Data Warehouse	https://www.gartner.com/imagesrv/media-products/pdf/solix/Solix-1-4WFTJQG.pdf
30	Microsoft (Microsoft, 2016)	Microsoft_Modern_Data_Warehouse_white_paper	www.microsoft.com
31	IBM (Difference, 2015)	Is the data warehouse dead?	tdwi.org/media

8.3 Relevantietabel

Hieronder worden alle hierboven geselecteerde publicaties beoordeeld op relevantie en aanwezige meetwaarden of succesfactoren

Nr.	Relevantie	Funcities / Factoren / Meetwaarden
1	Focus op data-project development met Agile (bijv Scrum) of planning tools (bijv Prince2, RUP). Wat werkt het beste? Kwantitatief onderzoek met 124 respondenten.	Factoren voor succes bij data-projecten: Mate van ondersteunen besluitvorming. Dit naast de bekende: tijd, kwaliteit, budget en scope / requirements. Overige factoren: - Agile waarden - Planning - Samenwerking IT en Business - Topmanagement ondersteuning - Technologie capaciteit - Complexiteit Conclusie: Voor data-projecten die sterk op een planning leunen, is top-management support de belangrijkste factor. Voor Agile-development is technologie de belangrijkste
2	Alleen een introductietekst, waarin verwezen wordt naar een onderzoek naar 2 organisaties die het succes en de failures van DWH presenteren	Het succes van DWH-implementaties gaat om meer focus op socio-technische factoren. Helaas niet bruikbaar binnen dit onderzoek
3	Wederom gaat het over het succes van een DWH-implementatie of -project, gebaseerd op algemeen IT-implementatie-succes. Ditmaal met de focus op ontwikkelingslanden zoals Jordanië.	Model bestaat uit 4 lagen: 1. DWH input door Systeemkwaliteit en kwaliteit van support 2. DWH gebruik, gemeten door waargenomen voordeel 3. DWH met BI: Data Management op individueel en corporate niveau 4. DWH succes door het behalen van strategische doelen
4	Algemeen artikel over data warehouse definities en functies met veel referenties uit het verleden. Paragraaf 4 beschrijft de uitdagingen.	Een groot DWH-project bij de FBI mislukte in 2005 door de complexe organisatiestructuur en communicatielijnen. Verwijzing naar (Goldstein, 2005). Succesfactoren: 1. Economisch succes: informatie voor de juiste personen om business te verbeteren 2. Politiek succes: Bewustwording en adoptie realiseren 3. Technologisch succes: Juiste tools DWH-implementaties mislukken vaak, door slecht support vanuit het management en inadequate user-betrokkenheid, wat samengevat kan worden in organisatorische factoren. Een ander onderzoek waaraan gerefereerd wordt (Abu Ali, 2010) is van 2 organisaties in de Verenigde Staten: Organisatorische factoren, omgevingsfactoren, projectfactoren, technische factoren en educatiefactoren.

5	Onderzoek naar volwassenheid van data warehouse processen (DWP) door vooraanstaande wetenschappers in dit vakgebied. Onderzoek onder 20 executives van 13 organisaties en een panel van experts. DWP is het data-productie-proces, inclusief business requirements, data design, architectuur, data mapping, ETL, datakwaliteit, etc. Het gaat dus over de ontwikkeling en beheer van een data warehouse! Net als CMM (Capab. Matur. Model)	DWH kent veel problemen: slechte performance, beschikbaarheid, complexe tools, slechte datakwaliteit en ontevreden gebruikers. DWP heeft net als CMM(I) 5 levels van volwassenheid: Zie bijlage 8.6 voor de Key Process Areas. 5.1 Metadata change management 5.2 DW technology change management 5.3 Defect prevention 5.4 DWP improvement program 5.5 Data governance 4.1 Data quality management 4.2 Quantitative process management 4.3 Integrated metadata quality management 4.5 Service level management 4.6 Financial services management 3.1 DWP definition 3.2 Organization process focus 3.3 Business metaData Management 3.4 Recovery management 3.5 Stakeholder management process 3.8 Inter-group coordination 3.9 Data quality assurance 3.12 Information delivery management 3.13 Service level agreement 3.14 DW governance 3.15 Operations/service quality assurance 3.16 Integrated operations/service mgmt. 3.17 Resource management 3.18 Configuration management 3.19 Training program 2.1 DW sponsor assurance 2.5 DW staffing 2.9 Subcontract management 2.10 Issue tracking 2.11 Standards setting 1.0 DW operations/service management done, but it is as good as the operations/service manager. In het panel kwamen de volgende succesfactoren aan de orde: - Voorspelbaarheid van DWH-projecten - Betere data analyse - Goede documentatie - Vertrouwen in de data(kwaliteit) - Transparante ROI - Waarneembaar verbeterde processen Geadviseerd wordt om meer onderzoek te doen in de vorm van veldwerk en enquêtes
6	Dit onderzoek gaat over de selectie van een data warehouse architectuur. Alle vormen komen	Welke criteria gelden er om tot een bepaalde architectuur te komen: - Informatieafhankelijkheid binnen een organisatie

	één voor één aan bod	- Urgentie - Routinematige taken / processen - Strategische kijk op data - Beschikbaarheid van kundige resources - Enthousiast IT-personeel kan management overtuigen - Champion rol om sponsoren te vinden
7	Indonesisch onderzoek wat zich richt op data warehouse implementaties. Het artikel bespreekt oude wijn in nieuwe zakken. De succesfactoren komen rechtstreeks van Watson	De factoren die genoemd worden, zijn: - Bureaucratie: Commitment en betrokkenheid van het management - User experience: Kennis van de business, techniek en de data - Development methode: Kimball of Inmon - Data warehouse schema: kies een architectuur en model dat bij het business proces past - Meten van (systeem) functionaliteit: Analyse, security, performance, data beschikbaarheid, datakwaliteit en data-integriteit - Eindgebruikerstools: Rapportage, Query, OLAP, etc
10	Het begrip “contextualized warehouse” wordt uitgelegd. Het is een combinatie van gestructureerde data en (ongestructureerde) documenten	Er worden geen succesfactoren beschreven
11	Er moet betaald worden voor dit artikel	
12	Het artikel focust zich op data warehouse adoptie en welke factoren hierbij spelen. Er wordt een bestaand model vanuit de IT toegepast op Decision Support Systems (DSS), waar een data warehouse ook onder valt. Het onderzoek is erg uitgebreid en biedt een mooie basis voor de onderzoeksvraag. Er zijn 7 factoren gevonden (5 organisatorische en 2 technische) Restrictie: Adoptie is een belangrijke succesfactor van een data warehouse, maar niet de enige. Data warehouse wordt gezien als een innovatieve en infrastructurele oplossing, maar daarmee wordt de software / tooling vergeten	De 7 factoren / meetwaarden die bepalend zijn voor data warehouse adoptie, zijn: 1. Organisatie commitment 2. Absorptievermogen 3. Organisatiegrootte 4. Organisatie scope van het data warehouse 5. Organisatie data omgeving 6. data warehouse voordelen 7. data warehouse complexiteit Aangegeven wordt dat vervolgstudie noodzakelijk is: data warehouse flexibiliteit, response-snelheid, standaardisatie, consistentie, integratie, beheer en datakwaliteit
13	Er worden 4 thema's beschreven rond data warehouse: Integratie,	De functie van een data warehouse is: Het ondersteunen van alle niveaus van management besluitvorming door het

	Implementatie, Intelligentie en Innovatie. Ook wordt beschreven hoe de gelaagde data warehouse-architectuur er uit moet zien, met per laag de functie. Veel tekst wordt besteed aan de architectuur en de functies van de gelaagde opbouw	ophalen, integreren, transformeren en interpreteren van zowel interne als externe data
14	Xu en Hwang geven in 2008 aan dat er nog niet veel structureel academisch onderzoek heeft plaatsgevonden op het gebied van data warehouse. Ze kijken of de succesfactoren van ICT-implementaties ook toepasbaar zijn op data warehouse en vinden 9 factoren. Daarna maken ze een model over hoe deze factoren met elkaar samenhangen	Er worden 9 factoren genoemd die een succesvolle implementatie garanderen: Operational, Technical, Schedule, Economic, System quality, Information quality, Individual benefits, Organizational benefits. Daarna maken ze een model over hoe deze factoren met elkaar samenhangen. Er wordt aangegeven dat er geen rekening is gehouden met de volgende meetwaarden: - Beheersbaarheid, gemak in onderhoud - Klant- of medewerkerstevredenheid - Data Management - Beschikbare alternatieven In de conclusie wordt aangegeven dat er meer onderzoek nodig is wanneer het aankomt op beheer en groei. "Easy to Manage" kan dan een nieuwe succesfactor zijn
15	Inmon mag niet ontbreken als ontdekker van het data warehouse. Vrijwel alle andere artikelen bevatten zijn definitie van het data warehouse. Dit artikel uit 1996 gaat over analyse (OLAP, Data Mining) en de noodzaak van een data warehouse hiervoor.	Een data warehouse levert per definitie: • Geïntegreerde data • Gedetailleerde en geaggregeerde data • Historische data • Metadata
16	Scott beschrijft de implementatie en beheer van hun data warehouse bij MIT. Hij beschrijft de functies en voordelen van een data warehouse, maar ook de valkuilen die vermeden moeten worden Ook al is het een artikel van het MIT, het is niet erg wetenschappelijk benaderd, maar meer een opsomming van de lessons learned.	De belangrijkste functie die in dit artikel van MIT beschreven wordt is: Het ondersteunen van makkelijke, flexibele en geïntegreerde rapportage. Naast de bekende functies en voordelen van een data warehouse, zijn er de volgende bevindingen: • Het herkennen en verbeteren van datakwaliteit in de organisatie • Het herkennen en verbeteren van data security procedures • Het spreken van één gemeenschappelijke business taal (metadata) De lessons learned uit dit artikel, kun je vertalen als succesfactoren. Deze zijn: • Herstructureer de data voordat het geïntegreerd en gemodelleerd wordt

		• Betrek de business tijdig en manage de verwachtingen • Betrek de stakeholders van de aanleverende bronsystemen tijdig en geef ze ruimschoots aandacht, omdat ze van cruciaal belang zijn
17	Artikel nog niet gevonden, maar online opgevraagd bij CBS	
18	Artikel nog niet gevonden, maar online opgevraagd bij CBS	
19	Artikel nog niet gevonden, maar online opgevraagd bij CBS	
20	Dit artikel focust op effectieve governance als belangrijkste succesfactor voor een data warehouse. Uit een eerdere studie van Watson zijn 8 data warehouse failures onderzocht en daaruit kwam Governance naar voren als belangrijkste organisatorische factor. Verder beschrijft dit artikel veel business benefits en kostenbesparingen door het data warehouse	Belangrijkste voordeel van een data warehouse is de "single view of the truth". Andere voordelen: • Consensus in data definities • Identificatie van de golden source • Betere data accuraatheid • Juiste en consistente rapportage Data warehouse governance is een noodzakelijke factor voor data warehouse succes. In dit artikel van Watson beschrijft hij een aantal aspecten: • Betrokkenheid van senior management • Creeer stuugroepen met de goede samenstelling, kennis, kunde en autoriteit en train ze in data warehouse • Bepaal Business behoefte voor de hele organisatie • Align data warehouse met organisatie doelen • Bepaal informatiebehoefte per user, afdeling • Stel juiste prioriteiten en verbeter het data warehouse stapje voor stapje (incrementeel) • Creeer datamodellen en definities • Beheer datakwaliteit • Align data warehouse architectuur en governance
21	Data warehouse-omgevingen en -projecten zijn belangrijk, maar ook risicovol, qua investering, tijd en inspanning. Volgens Al-Debei zijn er 2 categorieën issues & uitdagingen: Organisatorische & Technologische Dit artikel gaat vooral over de issues en uitdagingen bij de implementatie van een data warehouse. Dit is een prima uitgangspunt, maar er moet ook	Volgens Al-Debei zijn er 2 categorieën issues & uitdagingen die je kunt zien als succesfactoren: 1. Organisatie & aansturing: a. management commitment en support; b. project champion; c. user involvement en participatie; d. team skills en samenstelling, 2. Technologie: a. de selectie van een data warehouse architectuur; b. de creatie van een enterprise schema; c. data integratie en schaalbaarheid; d. data kwaliteit;

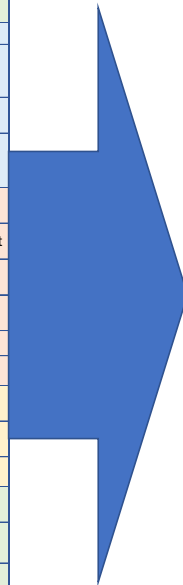
	gekeken worden naar (langdurig) operationele data warehouse-omgevingen	e. het design van human-computer interface of User Experience; f. het minen van een data warehouse; g. security en privacy risico's; h. netwerken en telecommunicatie
22	Burtescu (Emil & Claudia) zien de noodzaak van een data warehouse (en Data Mining) om belangrijke management-beslissingen te ondersteunen. Een data warehouse bevat een grote variëteit van data om een coherent beeld te geven van de business gesteldheid, op elk moment in de tijd Het begrip Data Mining wordt in dit artikel ook als Business Intelligence (BI) aangeduid	Een data warehouse heeft de volgende functionaliteit: 1. Het focust zich op de dagelijkse operatie en transacties 2. Het logt data en gooit nooit iets weg 3. Het verzamelt de data vanuit verschillende bronnen 4. Het integreert data
23	Dit artikel geeft aan dat veel data warehouse-omgevingen in de praktijk niet voldoende gebruikt worden. De oplossing hiervoor: semantische kubussen (SCM) op basis van Object Oriëntatie (OO) Vrij technisch onderzoek en oplossing naar semantische data kubussen	De functie van een data warehouse is om operationele systemen effectief te integreren in een omgeving die de strategie van een organisatie ondersteunt en de productiviteit van de besluitvoerders verbetert, door het consolideren, converteren, transformeren en integreren van operationele data. In de praktijk gebeurt dit onvoldoende. Door gebrek aan flexibiliteit, efficiëntie en schaalbaarheid. Een data warehouse groeit voortdurend (overloaded), waardoor het problemen geeft in het beheer en data-analyse (performance).
24	Watson claimde in 2001 als eerste & enige een wetenschappelijk onderzoek gedaan te hebben naar de succesfactoren van een data warehouse implementatie Dit onderzoek begint bij de basis: IT-implementatie succesfactoren en specialiseert zich daarna in data warehouse. Watson stelt: Er is geen onderzoek gedaan naar bestaande operationele data warehouse-omgevingen	Een doel van het data warehouse is om data te integreren vanuit de hele organisatie. Er zijn 3 factoren gevonden die doorslaggevend zijn voor het succes van een data warehouse: • Datakwaliteit • Systeemkwaliteit • Business behaalde voordelen Bij een data warehouse-implementatie zijn de volgende onderdelen van belang: • Organisatie • Project • Techniek En om dit te bereiken gelden de volgende succesfactoren: • Management support • Champion rol die het project overziet en dagelijks betrokken is • Resources, zoals geld, mensen, tijd • User participatie en betrokkenheid • Team skills

		• Bron systeem betrokkenheid en data standaardisatie • Development tools, hardware, software en methoden
25	Focus op organisatiestructuren en -regels rond een data warehouse bij grote (financiële) dienstverleners. Een data warehouse is data-georiënteerd, terwijl een organisatie proces-georiënteerd is	De organisatorische uitdagingen bij een data warehouse project zijn: - Organisatiedoelen alignen met project - Duidelijke verantwoordelijkheden voor data - Datakwaliteit management - Integratiemanagement op business processen - Multi-lagen-structuur, omdat een grote organisatie meerdere operationele systemen en Decision Support Systems (DDS) kent - Duurzaamheid, omdat data warehousing een permanent proces is Data-eigenaarschap moet geregeld worden op 3 niveaus: Inhoud, methode / bewerking en levering. Een data warehouse ondersteunt business processen, management processen en technische activiteiten.
26	Er worden 4 thema's beschreven rond data warehouse: Integratie, Implementatie, Intelligentie en Innovatie. Ook wordt beschreven hoe de gelaagde data warehouse-architectuur er uit moet zien, met per laag de functie. Het artikel focust op de strategische besluitvorming en de evaluatie ervan	De functie van een data warehouse is: Het ondersteunen van alle niveaus van management besluitvorming door het ophalen, integreren, transformeren en interpreteren van zowel interne als externe data
27	DWH/BI projecten zijn vaak te groot om Agile methodes toe te passen. Welke factoren beïnvloeden dit? 6 interviews zijn de basis van dit onderzoek. Half artikel: Alleen de introductie en conclusie is (gratis) gepubliceerd	Sommige delen van DWH-development zijn Agile toe te passen. Maar conclusie is dat Agile alleen in combinatie met klassiek projectmanagement is toe te passen
28	De onderzoekers willen, middels een literatuurstudie, aantonen dat het succes van een data warehouse niet alleen afhangt van datakwaliteit en systeemkwaliteit, maar ook van de relatiekwaliteit van de onderlinge data warehouse betrokkenen. De relatiekwaliteit is hoog als alle partijen dezelfde behoeften en verwachtingen hebben, door wederzijdse inzet,	Succes kent vele vormen. Er worden er 3 genoemd: - Economisch: positieve impact op de business - Politiek: kennis en kunde om gebruik te maken van het data warehouse - Technisch: toe te passen technische vaardigheid Ook de voordelen worden genoemd als succesfactor: - Beter besluitvorming - Beter klanttevredenheid - Snelle response op interne en externe events - Kostenbesparing - Verbeteren productiviteit - Concurrentievoordeel

	samenwerking en coördinatie. Er wordt verwezen naar bekende namen zoals Hwang, Wixom en Watson.	- Verbeten datakwaliteit Conclusie: Er is een verband tussen de relatiekwaliteit (met de factoren: Betrokkenheid, vertrouwen, communicatie, samenwerking en coördinatie) en data warehouse succes
29	Organisaties hebben een uitdaging tegenwoordig: ze weten niet welke grote volumes aan data ze hebben en wat ze ermee moeten doen. Een Logical Data Warehouse (LDW) is de oplossing om grote volumes aan data te verwerken voor zowel standaard rapportage op gestructureerde data, als Data Science activiteiten op (deels) ongestructureerde data.	Geen succesfactoren gevonden
30	Data is een asset en kan nieuwe inzichten genereren. Een klassiek data warehouse was nooit gebouwd om de explosie aan data en datatypes op te vangen. Om real-time data en Big data te ondersteunen, wordt er gesproken over een hybride vorm: Hybride of Modern Data Warehousing (Microsoft, 2016).	De toegevoegde waarde van een (klassiek) data warehouse is om business waarde te genereren op basis van rapportage en historische analyses. Nieuwe trends: - Toenemende data volumes - Real-time data - New types of (source) data - Cloud en hybride vorm - Advanced analytics en Machine Learning Een klassiek data warehouse ondersteunt geen inductieve analyse en heeft een vooraf gedefinieerd datamodel. Het kan de business verandersnelheid niet (meer) aan
31	Een onderzoek onder 90 grote organisaties wereldwijd (maar voornamelijk in de Verenigde Staten) moet bepalen of het klassieke data warehouse dood is en wordt vervangen door Big Data alternatieven.	88% heeft een data warehouses en 63% meer dan één. De grootte varieert tussen de 1 en 50Tb. Kosten: 35% van de respondenten is tevreden over de kosten van het beheer en onderhoud van hun data warehouse, 26% neutraal en 27% is ontevreden. Business waarde: 53% is tevreden, 33% neutraal en 7% is ontevreden. Big Data adoptie: 64% vindt Big Data belangrijk voor hun organisatie, 21% is neutraal en 12% vindt het niet belangrijk. Big Data initiatief: 36% is hiermee bezig of heeft iets live staan. 15% wil volgend jaar iets gaan doen en 32% is nog niet bezig hiermee. Einde van data warehouse: 43% denkt dat de een data warehouse en Big Data omgeving complementair zijn. 45% denkt dat Big Data het deels over gaat nemen. 7% denkt dat een data warehouse nooit zal verdwijnen. Conclusie: Een data warehouse leeft! Samen met Data Management en Big Data Analytics moet het een manier vinden om samen te werken

8.4 Vragen enquête en interviews

Factor / Categorie	Nr	Meetwaarde / Beschrijving
Waargenomen voordelen	V1	Ondersteunen primaire, secundaire en innovatieve processen
	V2	Ondersteunen van strategische doelstelling
	V3	Verbeterde individuele productiviteit
	V4	Verbeterde besluitvorming op strategisch en tactisch niveau
	V5	Verbeterde besluitvorming op operationeel niveau
	V6	Ondersteunen van analyse-doeleinden
Systeemkwaliteit	S1	Makkelijk in het gebruik en gebruiksvriendelijk
	S2	Eenvoudig en efficiënt
	S5	Security en autorisatie is goed geregeld
	S6	Integratie van meerdere bronsystemen
	S7	Goede beschikbaarheid
Datakwaliteit	D1	Hoogwaardige datakwaliteit
	D2	Voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens
	D3	Toereikende actualiteit van data (laadfrequentie en verversingsgraad)
	D4	Toegang tot geïntegreerde en consistente data
Organisatorische factor	O1	Doelstellingen zijn bekend en voldoen aan de business behoeften
	O2	Management vanuit de business en IT is betrokken en ondersteunt
	O3	Business (eind)gebruikers zijn actief betrokken
	O4	Goede communicatie tussen IT en business
	O5	Resources zijn beschikbaar gesteld
	O6	Er is een visie en roadmap
Technische factor	T3	Goede ontwikkelomgeving
	T4	Kundig IT-personeel
	T6	Veel data en bronsystemen, waardoor technisch complex
Economische factor	E1	Voldoende funding
	E2	Voordelen zijn bekend en worden financieel gemeten
	E3	Grootte organisatie
Beheer	B1	Verbeterpunten in de vorm van een backlog zijn bekend en geregistreerd
	B2	Continue verbeteringen door Agile werkwijze
	B3	Wijzigingen worden goed uitgevoerd en gemanaged
	B4	Goede responsetijd door schaalbaarheid
	B5	Gestandaardiseerd, gedocumenteerd en onderhoudbaar
	B6	Goede datakwaliteit aanleverende bronsystemen
	B7	Data Management is ingericht



#	Factor / Categorie	Nr	Vraagstelling
	Waargenomen voordelen	V10	Het data warehouse ondersteunt diverse primaire, ondersteunende en innovatieve processen
		V15 + V14	Het data warehouse draagt (in)direct bij aan strategische doelstelling, zoals het verlagen van kosten, verbeteren omzet & winst of een betere concurrentiepositie. Het DWH draagt daarmee (in)direct bij aan verbeterde bedrijfsprocessen
		V11	De informatie draagt zorg voor verbeterde individuele productiviteit
		V12	De informatie ondersteunt verbeterde besluitvorming op strategisch en tactisch niveau (oa dashboards, geaggregeerd niveau)
		V13	De informatie ondersteunt verbeterde besluitvorming op operationeel niveau (lijstwerk, laag detailniveau)
		V16	De data uit het data warehouse wordt gebruikt voor analyse-doeleinden, zoals OLAP, Data-mining of Data Science
		V01	Het data warehouse is makkelijk in het gebruik, heeft een gebruiksvriendelijke user interface of rapportagetool(s)
		V02	Het data warehouse werkt eenvoudig en efficiënt, zodat snel de benodigde informatie op te halen is
	Systeemkwaliteit	V05	De security en autorisatie is goed geregeld
		V03	Het data warehouse integreert data uit meerdere bronsystemen
		V04	Het data warehouse is altijd beschikbaar
		V06	De datakwaliteit van de data in het data warehouse is hoogwaardig
		V08 + V09	Het data warehouse bevat voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens die (deels) nergens anders beschikbaar zijn
	Datakwaliteit	V07	De actualiteit (laadfrequentie en verversingsgraad) van de data in het data warehouse is toereikend
		V08 + V09	Het data warehouse bevat voldoende meetwaarden (facts), contexten (dimensies) en afgeleide gegevens die (deels) nergens anders beschikbaar zijn
		V17	De doelstellingen van het data warehouse zijn bekend en beschreven en geven invulling aan de (originele) business behoeften
		V18	Het (top)management vanuit de business en IT is betrokken en ondersteunt de data warehouse-oplossing op alle manieren
	Organisatorische factor	V19	De business (eind)gebruikers zijn actief betrokken bij de ontwikkeling en het beheer van het data warehouse.
		V20	Er is regelmatig overleg tussen IT en business en er zijn continue verbetervoorstellen
		Zie vragen bij technische factor (V27) en beheerfactor (V28)	
		V21	Er is binnen uw organisatie een visie en roadmap ontwikkeld voor het data-domein
		V23	De ontwikkelomgeving van het data warehouse is toereikend. De tools en databases zijn actueel en worden ondersteund door de leveranciers
	Technische factor	V24	De kennis en kunde van het IT-personeel is toereikend. Er wordt gebruik gemaakt van externe consultants zodra dat nodig is
		V03	Het data warehouse ontsluit veel bronsystemen en is (daardoor) technisch complex
		V25	Er voldoende financiële draagvlak (funding) voor het beheer en verbetering van de data warehouse-omgeving
	Economische factor	V26	Het data warehouse levert toegevoegde waarde voor de organisatie, de business voordelen zijn bekend en worden kwalitatief en/of kwantitatief gemeten
		Gesloten vraag: Hoe groot is uw organisatie?	
	Beheer	V27	Er is een roadmap en/of backlog met eisen en wensen voor de (nabije) toekomst
		V28	Het data warehouse wordt continue verbeterd door een agile werkwijze of door regelmatig IT-projecten uit te voeren
		V29	De wijzigingen op het data warehouse worden telkens tijdig opgeleverd, binnen budget, scope en kwaliteitseisen
		V30	Het data warehouse is schaalbaar vwb de hardware en databases waarop het draait
		V31	Het data warehouse is gestandaardiseerd, gedocumenteerd en eenvoudig te onderhouden
		V22 + V32	De bronsystemen leveren data van voldoende kwaliteit: juist, volledig en tijdig. Dit is contractueel vastgelegd
		V33	Data management is ingericht en er wordt gewerkt met tools rond metadata, master data en data profiling

Figuur 8.3 Vragen enquête op basis van samengesteld succesmodel en dataverzameltabel

8.5 Resultaten enquête

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.Het DWH is makkelijk in het gebruik, heeft een gebruiksvriendelijke user interface of rapportagetool(s)	Enigszins van toepassing	Van toepassing	Veel van toepassing	Niet van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Niet van toepassing	Enigszins van toepassing	Van toepassing
2.Het DWH werkt eenvoudig en efficient, zodat snel de benodigde informatie op te halen is	Van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Niet van toepassing	Enigszins van toepassing	Enigszins van toepassing
3.Het DWH integreert data uit meerdere bronsystemen	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Veel van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing	Helemaal van toepassing
4.Het DWH is altijd beschikbaar	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Veel van toepassing
5.De security en autorisatie is goed geregeld	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Niet van toepassing	Van toepassing	Veel van toepassing
6.De datakwaliteit van de data in het DWH is hoogwaardig	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Helemaal van toepassing	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Van toepassing	Van toepassing
7.De actualiteit (laadfrequentie en verversingsgraad) van de data in het DWH is toereikend	Enigszins van toepassing	Helemaal van toepassing	Helemaal van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing
8.Het DWH bevat voldoende meetwaarden (facts) en contexten (dimensions)	Enigszins van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing
9.Het DWH bevat voldoende afgeleide variabelen die (deels) nergens anders beschikbaar zijn	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing
10.De informatieve waarde is goed en ondersteunt diverse primaire, ondersteunende en innovatieve processen	Enigszins van toepassing	Veel van toepassing	Helemaal van toepassing	Enigszins van toepassing	Helemaal van toepassing	Veel van toepassing	Helemaal van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing
11.De informatie draagt zorg voor verbeterde individuele productiviteit	Van toepassing	Niet van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing
12.De informatie ondersteunt verbeterde besluitvorming op strategisch en tactisch niveau (oa dashboards, geaggregeerd niveau)	Van toepassing	Helemaal van toepassing	Veel van toepassing	Niet van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing
13.De informatie ondersteunt verbeterde besluitvorming op operationeel niveau (fijstwerk, laag detailniveau)	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing
14.Het DWH draagt daarmee (indirect bij aan verbeterde bedrijfsprocessen	Van toepassing	Van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing
15.Het DWH draagt daarmee (indirect bij aan strategische doelstelling, zoals het verlagen van kosten, verbeteren omzet & winst en/of een betere concurrentiepositie	Van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing
16.De data uit het data warehouse wordt gebruikt voor analyse-doelenden, zoals OLAP of Data Science	Van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Niet van toepassing	Enigszins van toepassing	Veel van toepassing
17.De doelstellingen van het DWH zijn bekend en beschreven en geven invulling aan de (originele) business behoeften	Van toepassing	Van toepassing	Veel van toepassing	Niet van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing
18.Het (top)management vanuit de business en IT is betrokken en ondersteunt de DWH-oplossing op alle manieren, bijv. in de vorm van projecten (mensen en middelen)	Van toepassing	Van toepassing	Enigszins van toepassing	Niet van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Helemaal van toepassing
19.De business (eind)gebruikers zijn actief betrokken bij de ontwikkeling en het beheer van het DWH	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing	Helemaal van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing
20.Er is regelmatig overleg tussen IT en business en er zijn continue verbetervoorstellen	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing
21.Er is binnen uw organisatie een visie en roadmap ontwikkeld voor het data-domein	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Helemaal van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Enigszins van toepassing	Van toepassing	Helemaal van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing
22.De bronsystemen leveren data van voldoende kwaliteit: juist, volledig en tijdig. Dit is contractueel vastgelegd	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Niet van toepassing	Veel van toepassing	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing
23.De technologie van het DWH is toereikend. De tools en databases zijn actueel en worden ondersteund door de leveranciers	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing
24.De kennis en kunde van het IT-personeel is toereikend. Er wordt gebruik gemaakt van externe consultants zodra dat nodig is	Veel van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Veel van toepassing
25.Er voldoende financiële draagvlak (funding) voor het beheer en verbeteren van de DWH-omgeving	Van toepassing	Veel van toepassing	Enigszins van toepassing	Helemaal van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing
26.Het DWH levert toegevoegde waarde voor de organisatie, de business voordelen zijn bekend en worden kwalitatief en/of kwantitatief gemeten	Van toepassing	Veel van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing
27.Er is een roadmap of backlog met eisen en wensen voor de (nabije) toekomst	Van toepassing	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Niet van toepassing	Veel van toepassing
28.Het DWH wordt continue verbeterd door een agile werkwijze of door regelmatige IT-projecten uit te voeren	Van toepassing	Veel van toepassing	Helemaal van toepassing	Niet van toepassing	Van toepassing	Van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Veel van toepassing	Van toepassing	Niet van toepassing	Helemaal van toepassing
29.De wijzigingen op het DWH worden telkens tijdig opgeleverd, binnen budget, scope en kwaliteitseisen	Enigszins van toepassing	Veel van toepassing	Veel van toepassing	Niet van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing
30.Het DWH is schaalbaar vwb de hardware en databases waarop het draait	Enigszins van toepassing	Veel van toepassing	Enigszins van toepassing	Niet van toepassing	Van toepassing	Enigszins van toepassing	Helemaal van toepassing	Veel van toepassing	Van toepassing	Van toepassing	Niet van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing
31.Het DWH is gestandaardiseerd, gedocumenteerd en eenvoudig te onderhouden	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing
32.De bronsystemen leveren data van voldoende kwaliteit: juist, volledig en tijdig. Dit is vastgelegd in een overeenkomst (SLA of DDA)	Van toepassing	Enigszins van toepassing	Van toepassing	Niet van toepassing	Veel van toepassing	Veel van toepassing	Veel van toepassing	Enigszins van toepassing	Enigszins van toepassing	Van toepassing	Van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing
33.Data management is ingericht en er wordt gewerkt met tools rond metadata, master data en data profiling	Van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Van toepassing	Enigszins van toepassing	Enigszins van toepassing	Enigszins van toepassing	Veel van toepassing

Figuur 8.4 Resultaten enquête

8.6 Data Warehouse Processen (DWP) en de Key Process Areas (KPA's)

Levels	KPAs for "Development"	KPAs for "Operations/Services"
Optimizing	Analyzes the defects to understand their causes and evaluates the process to prevent the known type of defects. Continuous improvement is also done. KPAs are: <u>5.1</u> Metadata change management <u>5.2</u> DW technology change management <u>5.3</u> Defect prevention <u>5.4</u> DWP improvement program	Analyzes the defects to understand their causes and evaluates the process to prevent the known type of defects. Continuous improvement is also done. KPAs are: <u>5.1</u> Metadata change management <u>5.2</u> DW technology change management <u>5.3</u> Defect prevention <u>5.4</u> DWP improvement program <u>5.5</u> Data governance
Managed	Quantitative and statistical techniques are used to manage process performance and quality of DW products under development. Focus on developing organization-wide process database to collect and analyze the process data. KPAs are: <u>4.1</u> Data quality management <u>4.2</u> Quantitative process management <u>4.3</u> Integrated metadata quality management <u>4.4</u> Data change management <u>4.5</u> Service level management	Quantitative and statistical techniques are used to manage process performance and quality of DW product in operation/service. Focus on developing organization-wide process database to collect and analyze the process data. KPAs are: <u>4.1</u> Data quality management <u>4.2</u> Quantitative process management <u>4.3</u> Integrated metadata quality management <u>4.5</u> Service level management <u>4.6</u> Financial services management
Defined	Project management is based on a defined process. Focus on process. KPAs are: <u>3.1</u> DWP definition <u>3.2</u> Organization process focus <u>3.3</u> Business metadata management <u>3.5</u> Stakeholder management process <u>3.6</u> Alignment of architecture <u>3.7</u> Integrated infrastructure management <u>3.8</u> Inter-group coordination <u>3.9</u> Data quality assurance <u>3.10</u> DW product engineering <u>3.11</u> Peer review <u>3.13</u> Service level agreement <u>3.17</u> Resource management <u>3.18</u> Configuration management <u>3.19</u> Training program	DW operations/service management is based on a defined process. Focus on process. KPAs are: <u>3.1</u> DWP definition <u>3.2</u> Organization process focus <u>3.3</u> Business metadata management <u>3.4</u> Recovery management <u>3.5</u> Stakeholder management process <u>3.8</u> Inter-group coordination <u>3.9</u> Data quality assurance <u>3.12</u> Information delivery management <u>3.13</u> Service level agreement <u>3.14</u> DW governance <u>3.15</u> Operations/service quality assurance <u>3.16</u> Integrated operations/service mgmt. <u>3.17</u> Resource management <u>3.18</u> Configuration management <u>3.19</u> Training program
Repeatable	Documented and realistic plans are the basis for managing project. Project Management is in place. KPAs are: <u>2.1</u> DW sponsor assurance <u>2.2</u> DW program planning <u>2.3</u> DW project planning <u>2.4</u> Business justification <u>2.5</u> DW staffing <u>2.6</u> Requirements management <u>2.7</u> Scope design and verification <u>2.8</u> Project tracking <u>2.9</u> Subcontract management <u>2.10</u> Issue tracking <u>2.11</u> Standards setting	Documented and realistic policies and procedures are the basis for managing operations. DW operations/service management is in place. KPAs are: <u>2.1</u> DW sponsor assurance <u>2.5</u> DW staffing <u>2.9</u> Subcontract management <u>2.10</u> Issue tracking <u>2.11</u> Standards setting
Initial	1.0 Project management done, but it is as good as the development project manager	1.0 DW operations/service management done, but it is as good as the operations/service manager

Figuur 8.5 Data warehouse maturity levels van Sen et al. (Sen A. R., 2012)