

Faalfactoren

bij data warehouses

Grundsätzlich IT

M: 06-41966000

W: www.grundsatzlich-it.nl

E: info@grundsatzlich-it.nl



R. Kunenborg

Versie 1.62 / 25 September 2009

Op dit werk is een [Creative Commons Licentie](https://creativecommons.org/licenses/by-nd/4.0/) van toepassing.



Contents

Faalfactoren bij data warehouse implementaties	3
Inleiding	3
Achtergrond.....	3
De impact van best practices	4
BI-projecten zijn IT-projecten.....	5
Specifieke data warehouse issues.....	6
Problemen met dataleveranciers	6
Problemen met eindgebruikers.....	6
Een data warehouse is een proces en geen project.....	7
Conclusie	7
Kader	8

Faalfactoren bij data warehouse implementaties

Inleiding

Binnen de business intelligence gemeenschap circuleren er vrij veel artikelen over het falen van data warehouse projecten. Meestal worden daar allerlei 'best practices' aan vastgeknoopt. In dit stuk wil ik een lans breken voor het idee dat data warehouse projecten grotendeels gelijk zijn aan normale IT-projecten, of in ieder geval minder uniek dan wel wordt geclaimd. Voor de enkele factoren die wel afwijken geef ik tips om de impact te verminderen waar projectleiders en architecten hun voordeel mee kunnen doen.

Achtergrond

Het mislukken van een fors percentage van alle data warehouse projecten vormt een behoorlijke bron van slapeloze nachten voor de CIO. Business Intelligence is namelijk cruciaal voor het overleven van bedrijven.

Met name het ontbreken van operationele business intelligence, waardoor de medewerkers alle informatie hebben om correcte beslissingen te kunnen nemen, is in een kredietcrisis erg gevaarlijk. Een concurrent die in snel veranderende tijden wel effectief op veranderende markten kan reageren, duwt je de markt uit. En voor bedrijven die dan in moeilijkheden geraken is het ontbreken van strategische business intelligence over hoe het bedrijf er voor staat, dodelijk. Deze bedrijven komen bij hun bank om hulp vragen en blijken er dan veel slechter voor te staan dan gedacht waardoor ze niet meer te redden zijn.

Business intelligence is dus cruciaal voor een bedrijf, maar vooral voor de *toekomstige* bedrijfsvoering. Het mislukken van een business intelligence traject verandert namelijk niets aan de *huidige* bedrijfsvoering. Dat maakt het moeilijk om een eenduidig beeld te krijgen van mislukkende business intelligence trajecten, want wanneer is een project nu precies mislukt?

Een falende invoering van een nieuw ERP-systeem is goed te zien, omdat de klanten hun spullen niet of niet op tijd krijgen, of met verkeerde rekeningen worden opgezadeld. Maar bij een falend business intelligence systeem kan falen variëren van het stopzetten van het project en afschrijven van het uitgegeven geld, tot iets subtiels als het niet of nauwelijks gebruiken van de mogelijkheden door de beoogde doelgroep.

Studies naar falende data warehouse projecten gebruiken dan ook verschillende definities voor falen. Daardoor lopen de schattingen voor de aantallen mislukte implementaties nogal uiteen: van 10-20% bij een studie uit 2005, tot 50% voor een studie van Gartner in datzelfde jaar en een jaar later worden zelfs problemen genoemd met tot 81% van alle data warehouse implementaties bij een studie van The Datawarehouse Institute (TDWI).

Het gaat om forse bedragen. Je zou verwachten dat met zo'n potentiële faalkans, data warehouse projecten slechts met grote aarzeling worden ondernomen. Dat is echter niet het geval. Dat komt omdat data warehouse projecten het niet slechter doen dan gewone projecten.

Het CHAOS rapport van de Standish Group uit 2003 analyseerde 13522 IT-projecten, en liet zien dat 15% van alle projecten 'failed' en 51% 'challenged' waren. Dat zijn cijfers die vergelijkbaar zijn met die van Hugh J Watson's onderzoek uit 2005¹, waarin op basis van een enquête onder 454 organisaties bleek dat ongeveer 12% van alle data warehouse initiatieven in de problemen zat, en zo'n 40-50% nog bezig was met de implementatie. De laatste groep ging mogelijk wel over het beschikbare budget of over de geplande tijd heen. Deze cijfers komen goed overeen met de cijfers van Gartner en andere onderzoeken², als we de definities gelijktrekken.

Opvallend is daarbij wel dat de cijfers gaan afwijken als er niet op basis van best practices wordt gewerkt. In het onderzoek van Hugh J Watson kwam naar voren dat bij projecten die op basis van 'federated' of 'independent datamart' architecturen waren opgezet, bijna 20% van de projecten in de problemen zat –twee keer zoveel als bij architecturen gebaseerd op Kimball's geconformeerde dimensies of Inmon's genormaliseerde centrale data warehouse. Daarom is een betere conclusie: data warehouse projecten mislukken even vaak als andere IT-projecten, als beide projecten op basis van 'best practices' zijn opgezet. Laten we daarom eens kijken naar die 'best practices', welke er relevant zijn en wanneer.

De impact van best practices

Een onderzoek naar 'best practices' gebaseerd op ervaring met mislukte data warehouse projecten is lastig. Er zijn relatief weinig harde feiten te vinden over mislukte data warehouse projecten. Een onderzoek uit 2002³ in meer dan 5 zoekmachines en 12 tijdschriften vond in de periode tussen 1997 en 2001 slechts 7 artikelen waarin mislukte data warehouse publicaties werden beschreven. Over dezelfde periode werden er 25 artikelen aangetroffen waarin beschreven werd hoe het dan wel moest. Die verhouding leek zich ook uit te strekken over andere periodes.

Voor dit artikel heb ik de zoektocht herhaald met behulp van Google. Over de periode 2001 tot en met 2009 vond ik slechts 5 artikelen met een beschrijving van een mislukking waar de schrijver zelf bij betrokken was. Daartegenover stonden 19 artikelen waarin werd beschreven hoe het dan wel moest. Aan de verhouding tussen artikelen over succesvolle en minder succesvolle implementaties is dus niet veel veranderd.

Een mogelijke verklaring voor het gebrek aan artikelen over mislukte projecten is de terechte huivering bij veel bedrijven om een dure mislukking openbaar te maken. Een mislukt IT-project heeft mogelijk gevolgen voor de aandelenkoers en doet de carrière van de verantwoordelijke manager ook weinig goed.

Het belangrijkste artikel over faalfactoren in de data warehouse omgeving blijft dan ook het veelgeciteerde 'The politics of data warehousing' van Marc Demarest uit 1997. Vooral de tegenmaatregelen blijven relevant en zijn waar van toepassing verwerkt in dit document.

Om een succesvol data warehouse project te voltooien is het echter niet genoeg om een rij 'best practices' te selecteren en die aan te houden. Uit een onderzoek uit 2007 door Hwang en Xu blijkt namelijk dat afhankelijk van het doel van een data warehouse project sommige adviezen veel meer

¹ Are Data Warehouses Prone to Failure?, Hugh J. Watson, 9 november 2005, Business Intelligence Journal

² Lessen uit ICT-projecten bij de overheid, 29 november 2007, Algemene Rekenkamer

³ Critical Factors for Data Warehouse Failure, Mark N. Frolick & Keith Lindsey, 2002, Business Intelligence Journal

impact hebben dan andere adviezen⁴. Zo is het voor een data warehouse wat is opgezet met als doel 'verbeterde concurrentiepositie' cruciaal om te letten op 'goed gedefinieerde zakelijke baten en eisen', 'project management' en 'meetbare baten voor het bedrijf' maar is de 'passende datawarehouse technologie' niet beslissend voor de uitkomst.

Verder zien we dat inzichten na verloop van tijd verschuiven. Sommige aandachtspunten blijven constant geldig, maar andere punten komen na het jaar 2000 veel minder voor in aanbevelingen. Dat geldt bijvoorbeeld voor de factor 'het project moet passen bij de bedrijfsstrategie en bedrijfsdoelstellingen'. Tegelijkertijd komt de factor 'zorg voor steun vanuit het management' juist vaker voor. Als het data warehouse niet duidelijk in lijn is met de bedrijfsstrategie, dan zul je inderdaad meer moeite moeten doen om het management te overtuigen. Het kan echter ook zijn dat men tegenwoordig veel meer rekening houdt met de bedrijfsstrategie.

Los van de moeilijkheden rond het bepalen van de reden waarom een 'best practice' werkt, zijn sommige 'best practices', zoals de 'best practices' die met de manier van leidinggeven te maken hebben, ook cultuurgebonden. Daarom is het riskant om zonder verdere analyse te vertrouwen op beschreven 'best practices' zonder verdere context.

BI-projecten zijn IT-projecten

Als we kijken naar de beschreven onderzoeken dan zien we dat de beschreven faalfactoren en 'best practices' niet specifiek gericht zijn op data warehouse projecten. Er zijn veel factoren die voor alle IT-projecten gelden:

- Zorg voor realistische verwachtingen bij de eindgebruikers;
- Let op de businesscase. Een ontbrekende businesscase, of een businesscase zonder duidelijke onderbouwing is een bron van problemen, maar een businesscase die voor elke afdeling afzonderlijk positief moet zijn is bij een datawarehouse vaak net zo min haalbaar als bij een ERP-implementatie;
- Zorg voor voldoende budget, geldgebrek nekt elk project;
- Zorg voor goed omschreven systeemeisen, ga niet bouwen op onduidelijke basis;
- Hou de oorspronkelijke doelstelling in de gaten – vermijd het sluipenderwijs uitbreiden van de eisen en wensen ('scope creep');
- Let op de 80/20 regel: 80% van de functionaliteit bouw je in 20% van de tijd (en geld), en andersom;
- De faalkans neemt exponentieel toe met de duur en omvang van het project;
- Betrek de eindgebruikers bij de oplossing, ga uit van een 50/50 tijdsverdeling tussen ontwikkelaars en eindgebruikers (vooral voor informatieanalyse en testen);
- Gebruik de goede technieken en zorg voor medewerkers met voldoende kennis, maar let er op dat de techniek dienend is en niet leidend;
- Zorg voor een architectuur die in lijn is met de bedrijfsdoelstellingen, zodat je weet wat voor soort systemen nodig zijn en je niet eindigt met meerdere gelijksoortige systemen⁵;

⁴ The Effect of Implementation Factors on Data Warehousing Success: An Exploratory Study, Mark I. Hwang & Hongjiang Xu, 'Journal of Information, Information Technology, and Organizations', Volume 2, 2007

⁵ Management Update: What to Expect When Implementing a Business Application Data Warehouse, Frank Buytendijk & Ted Friedman, Gartner Research, 18 mei 2005

Het interessante aan bovenstaande aandachtspunten is niet dat ze onbekend zijn, maar meer dat niet iedereen beseft dat ze ook voor data warehouse projecten onverkort gelden. Daar laat ik het even bij. Interessanter zijn namelijk de problemen die specifiek zijn aan data warehouse projecten, omdat we daar als data warehouse experts kennis kunnen bijdragen die een reguliere projectleider misschien niet heeft.

Specifieke data warehouse issues

Problemen met dataleveranciers

Het gebeurt natuurlijk regelmatig dat systemen data aan elkaar moeten leveren. Voor data warehouses is dit echter *altijd* een kwestie. Soms hebben we daarbij te maken met dataleveranciers die data niet, niet goed, of inconsistent aanleveren en daarvoor niet ter verantwoording worden geroepen.

Als het data warehouse wordt onderhouden door een afdeling die los staat van de afdeling die de operationele systemen onderhoud, zonder dat er een formele klant-leverancierverhouding is met sancties voor niet of slecht aanleveren van gegevens, dan wordt het makkelijk om bij wijzigingen te vergeten het data warehouse team in te lichten. Elke interface betekent extra werk en de IT-afdeling voor de operationele systemen is waarschijnlijk al druk genoeg met de eigen taakstelling.

Zorg daarom voor contracten met de dataleveranciers, ook al zijn ze intern. De aanlevering van data aan het data warehouse is cruciaal. Betrek hier het management bij door als eerste rapportage een rapportage op te leveren over de aanleveringen en de datakwaliteit. Dit zorgt zowel voor een duidelijk signaal naar de leveranciers dat er op ze wordt gelet, als naar het management dat er voortgang wordt geboekt.

Voorkom in ieder geval altijd dat het data warehouse team verantwoordelijk wordt voor het zelf ophalen van gegevens. Zodra dat gebeurt ben je de controle kwijt en dit leidt op de langere termijn tot het falen van het data warehouse.

Problemen met eindgebruikers

Problemen met eindgebruikers zijn grotendeels te verklaren uit het overschrijden van 'verdragsgrenzen'. Binnen elke organisatie bestaan koninkrijkjes gebaseerd op het eigendom en beheer van kennis. Een data warehouse project breekt al snel door de bestaande grenzen heen en zorgt daarmee voor een verstoring van de orde door het openbaar maken van problemen met de datakwaliteit in systemen en het lastig of onmogelijk maken van manipuleren van cijfers.

Het bouwen van een data warehouse is een politiek proces dat vrijwel alle onderdelen van de organisatie zal raken. In die onderdelen bevinden zich ook mensen die het niet zien zitten om mee te werken, hun data ter beschikking te stellen of om ter verantwoording te worden geroepen voor het manipuleren van de gegevens in hun voordeel. Die mensen zijn soms ook machtiger dan de IT-afdeling.

Ga daarom pas een data warehouse bouwen als het een pijnpunt oplost voor eindgebruikers. Als ze niet schreeuwen om een oplossing, dan heeft het weinig zin iets te bouwen. Ook al omdat de gebruikers meestal niet gedwongen kunnen worden hun vertrouwde 'spreadmarts' los te laten, zal je

tijdens en na de implementatie moeten laten zien dat het data warehouse voor de pijnpunten een echte oplossing biedt. Robuuste en betrouwbare datalevering draagt hier natuurlijk stevig aan bij.

Een marketing- en trainingsplan om de voordelen van de oplossing te demonstreren is eveneens nuttig. Betrek het management van de eindgebruikers bij de oplossing. Zorg dat het hun project wordt. Een combinatie van eindgebruikers die de oplossing willen hebben en hun management dat hen daarbij steunt is vaak effectief in het afbreken van interne weerstand. Steun van het management zonder steun van de eindgebruikers levert weinig op.

Een data warehouse is een proces en geen project

Omdat kenniswerkers met hulp van het data warehouse opeens over de traditionele informatiegrenzen kunnen kijken, zullen ze als alles goed gaat meer rapporten vragen dan vroeger. Langzaam verschuift hun werk van puur rapporteren, naar analyseren. Maar als cruciale projectmedewerkers al weer zijn ingezet op andere projecten dan kunnen de gevraagde extra rapporten niet worden gebouwd. De vervolgens optredende vertraging leidt tot teleurstelling bij de eindgebruikers en zorgt dat een succesvol project alsnog mislukt. Hou dus rekening met een toename van de vraag naar ondersteuning van de eindgebruikers en zorg daarin voor continuïteit.

Het data warehouse beweegt mee met het bedrijf. Dit komt voort uit het grensoverschrijdende karakter van bedrijfsinformatie. Daarom zal bij elke wijziging in de bedrijfsvoering die verder gaat dan een cosmetische aanpassing het data warehouse ook moeten worden bijgewerkt. Als dat niet wordt meebegroet in de kosten van de wijziging van het operationele systeem gaat het data warehouse steeds verder achterlopen bij de rest van het bedrijf. Zorg dus voor een wijzigingsbudget.

Een succesvol data warehouse kent ook uitbreidingen. Dat is in het bijzonder in organisaties die elke afdeling afzonderlijk de kosten laten dragen voor hun projecten, lastig om te bekostigen. Stel namelijk dat een nieuw systeem moet worden ontsloten: voor elke losse afdeling kan het best zo zijn dat de kosten niet opwegen tegen de baten. Maar een data warehouse is typisch een systeem waarbij de waarde van de informatie toeneemt met het aantal verbindingen wat je tussen de gegevens kan leggen. Het kan dus goed zijn dat pas bij ontsluiten van het vierde systeem de baten boven de kosten uit gaan komen, of dat pas als de gegevens zijn geladen de baten duidelijk worden. Zorg dus voor een perspectief op langere termijn dan alleen de eerste datamart.

Richt dus een organisatie in rond het data warehouse, een Business Intelligence Competence Center (BICC), die het kan blijven uitbreiden, promoten, de benodigde ondersteuning kan bieden aan eindgebruikers en zo zorgt dat het data warehouse in leven blijft. Zorg dat een data warehouse waarde toevoegt aan het bedrijf is een proces, geen project.

Conclusie

Een data warehouse project is een IT-project zoals alle andere. Wie de basale zaken daarin niet goed oppakt, krijgt dan ook de standaard problemen van niet goed verlopende IT-projecten. Maar buiten dat zijn er nog enkele extra factoren waar op moet worden gelet. Het is de toegevoegde waarde van de business intelligence expert om die factoren in een project te onderkennen en te voorkomen.

Kader

	Goed gedefinieerde zakelijke noodzaak/baten	Top management ondersteuning	Betrokkenheid en deelname van eindgebruikers	Datakwaliteit van de bron	Proper development technology	Adekwate IS staf en consultants	Project management (teamwork)	Practisch implementatie-schema	Goede planning/scoping van het project	Genoeg budget	Meetbare zakelijke baten
Makkelijk in het gebruik	X		X	X				X		X	
Informatie is snel op te vragen			X		X					X	
Meer informatie beschikbaar							X			X	
Betere informatie-kwaliteit				X			X				
Verbeterde productiviteit			X			X				X	X
Betere beslissingen	X					X	X				X
Verbeterde bedrijfsprocessen										X	X
Verbeterde concurrentiepositie	X						X				X

Correlaties tussen data warehouse succesfactoren (verticaal) en implementatie-factoren (horizontaal).

Samengevat uit 'The Effect of Implementation Factors on Data Warehousing Success: An Exploratory Study', Mark I. Hwang & Hongjiang Xu, 'Journal of Information, Information Technology, and Organizations', Volume 2, 2007.