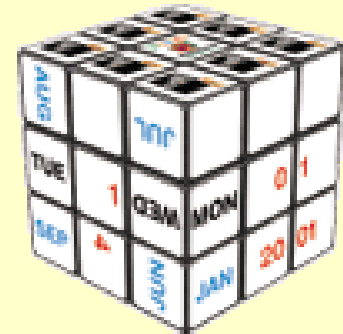
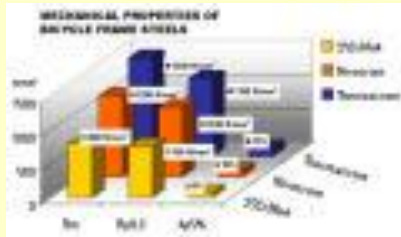


Data Warehouse Testen

19 november 2007 – Vereniging TestNet – Thema-avond DWH Testen

De 8 fouten die gemaakt worden Traditioneel testen versus DWH/BI testen



Anita van Berghenegouwen
business intelligence consultant bij Centennium BI expertisehuis



*Alle kennis, kunde en ervaring die nodig is om te beantwoorden aan complexe business intelligence en datawarehousing vraagstukken **onder één dak...***

Sinds 1998 houdt Centennium BI expertisehuis zich bezig met het ondersteunen van haar opdrachtgevers bij het beantwoorden van complexe informatiserings vraagstukken.

Kort samengevat kunt u bij ons terecht voor **advies** (ten aanzien van alle strategische, programmatische, procesmatige en technische aspecten van business intelligence, datawarehousing, performance management en informatiemanagement), **projecten** (uitvoering van of participatie in uw BI projecten), **capaciteit** (tijdelijk inzetbare BI expertise op basis van detachering) en **opleidingen** (opleidingen en trainingen voor gebruikers, ontwikkelaars en beheerders van uw BI oplossingen).

Centennium is een van de weinige daadwerkelijk onafhankelijke leveranciers van BI expertise in Nederland. Met ruim 40 zeer ervaren en veelal gecertificeerde business intelligence consultants en specialisten bedient Centennium een gevarieerde groep opdrachtgevers. Bijvoorbeeld binnen de overheid, bij financiële dienstverleners, industrie en de top van het MKB.

Anita van Berghenegouwen

business intelligence consultant

a.vanberghenegouwen@centennium.nl



Anita heeft een bedrijfskundige HBO opleiding genoten en verwacht in de loop van 2008 haar titel Master of Science (business processes en ICT) af te ronden. Door trainingen, ervaring en een bovenmatige interesse in het vakgebied heeft zij zich ontwikkeld tot een uiterst breed inzetbare business intelligence consultant.

In het recente verleden vervulde zij als business intelligence consultant diverse rollen, waaronder die van testcoördinator, projectleider en ontwerper. Rollen waarin zij een waardevolle bijdrage heeft geleverd aan de BI/DWH successen van onder andere Dienst Justitiele Inrichtingen, KPN, Imperial Tobacco, TPG post en het Ministerie van Buitenlandse Zaken.

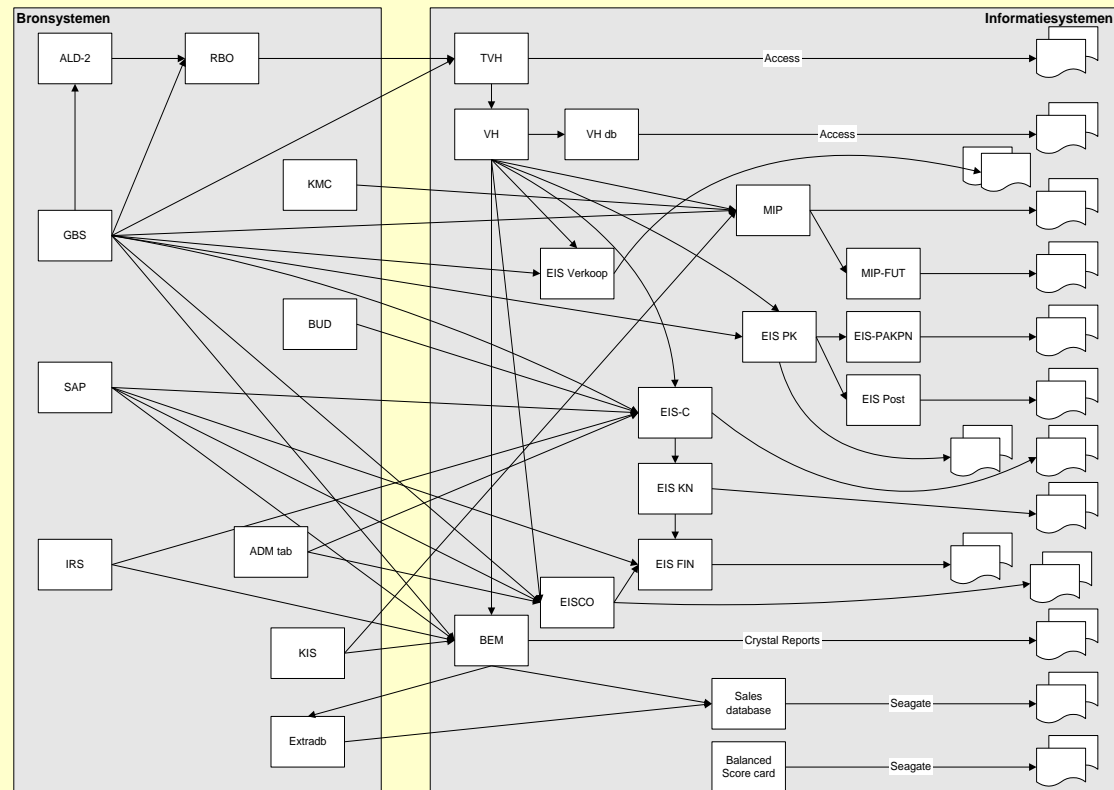
Door haar ervaring met testtrajecten binnen zowel "gewone" ontwikkel projecten als business intelligence en datawarehousing projecten kan Anita duidelijk toelichten waar er belangrijke verschillen in het testproces optreden. Ook kan zij als geen ander de DO's en DON'T's bij testen in BI/DWH omgevingen onder woorden brengen.

Onderwerpen

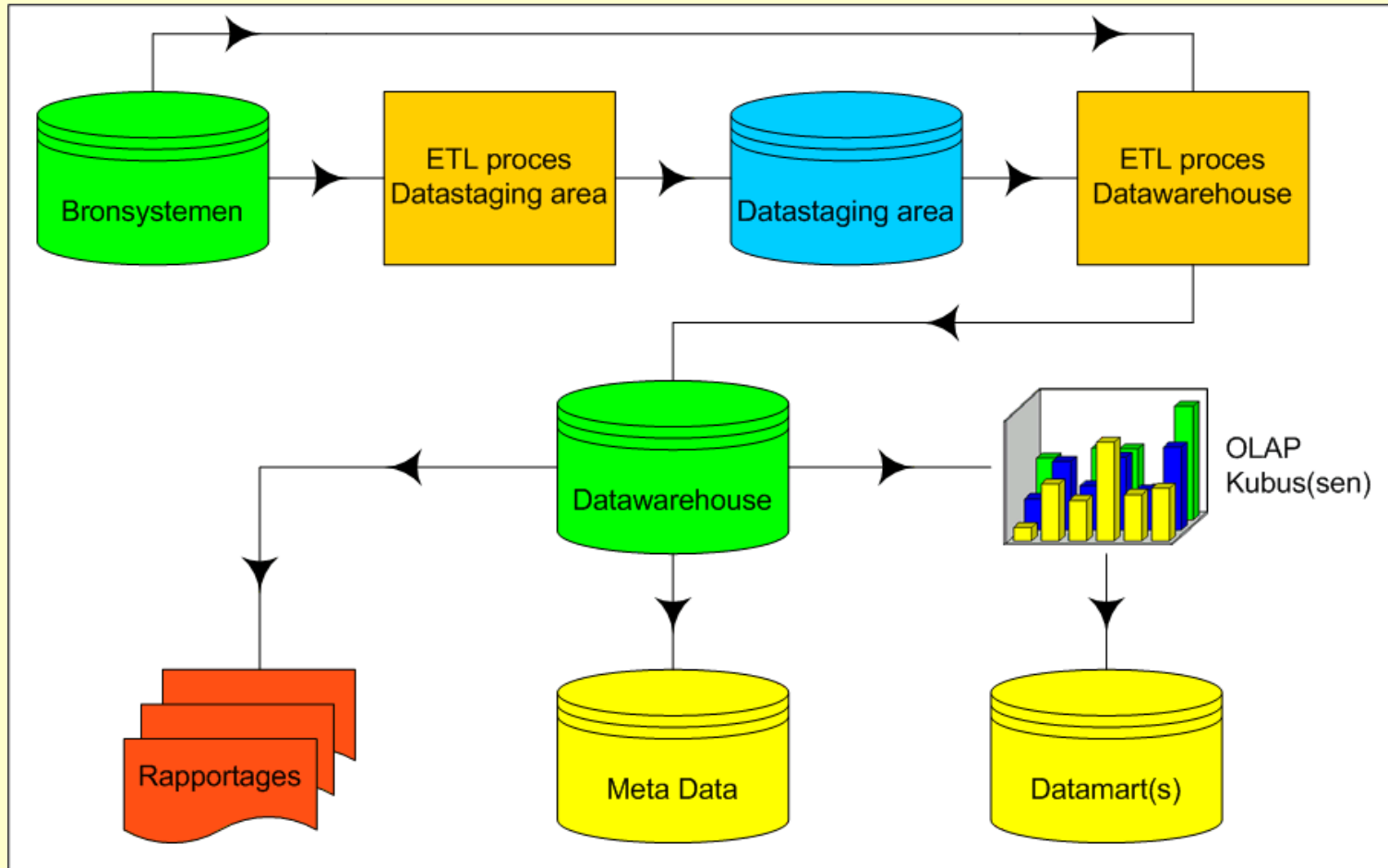
1. Data Warehouse (DWH)
2. De 8 fouten
3. Traditioneel testen versus DWH/BI testen
4. Typisch testproces

Waarom Data Warehousing

Veel organisaties concluderen dat de besturingsinformatie niet van voldoende kwaliteit is en besluiten over te gaan tot een Data Warehouse



Technische Architectuur

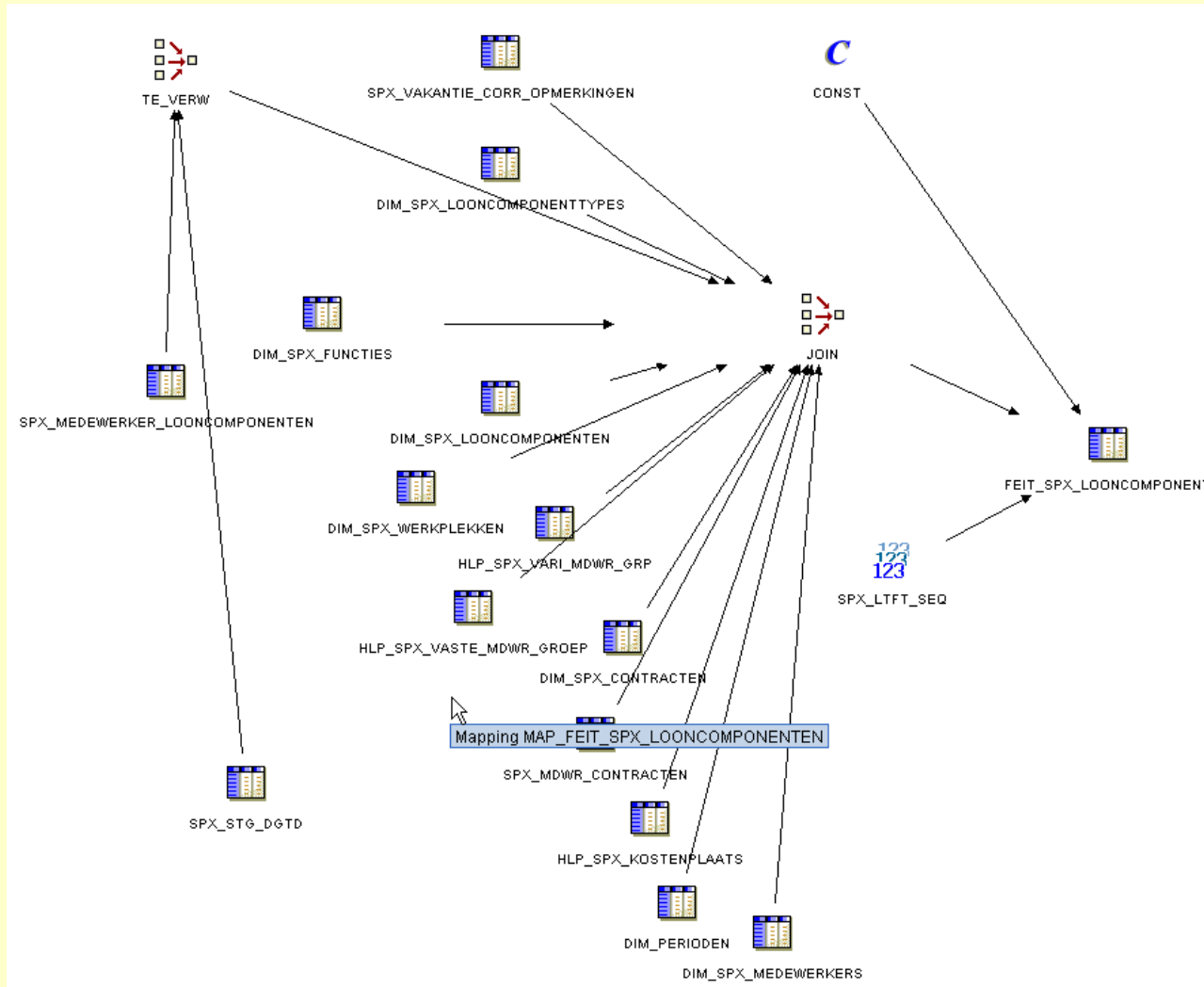


De 8 fouten

Fout 1: De gebruikers niet betrekken

- Eindgebruikers zijn de reden dat een Data Warehouse bestaat
- Falen → Waarom? De gebruikelijke redenen

Fout 2: Rapporten testen – geen data



Fout 2: Rapporten testen – geen data

- Alleen het testen van de rapportages zal niet genoeg zijn
- Testen van het Extraheren, Transformeren en Laden (ETL) vraagt prioriteit

Fout 3: Nalaten van vergelijking data DWH met data uit het bronsysteem

- Eindgebruikers zullen de data in het data warehouse niet 'geloven' als zij niet zeker weten dat deze data accuraat is
- Reserveer tijd om rapporten te ontwerpen die de data warehouse rapporten in verband kunnen brengen met de rapporten uit de legacy systemen

Fout 4: Onderwaardering voor het testwerk

- Door gebruik van ETL tools, minder geschreven code voor de test fase
- De hoeveelheid code kan het aantal test issues niet voorspellen

Fout 5: Niet definiëren van heldere verwachtingen

- Testen op data validatie kan heel grondig zijn maar het zal nooit ieder data issue dat zich voordoet kunnen adresseren
- Zorg ervoor dat testers kennis hebben van de kwaliteit van de data in het data warehouse
- Draag er zorg voor dat verwachtingen zijn gecommuniceerd naar gebruikers, ver voordat het testen start

Fout 6: Onderschatten data kwaliteit en data integratie

- Zorg ervoor dat er een bepaald niveau van tevredenheid is met de kwaliteit van de data voordat gestart wordt met het uitvoeren van ETL processen
- Als het dagelijks proces van laden 25 uur duurt, zal men terug moeten naar het ontwerp
- Verrassingen zullen het vertrouwen in het systeem doen vervagen
- Het onderschatten van het ETL proces

Fout 7: Kiezen van de verkeerde testers

- Gebruik testers die enige kennis hebben van DWH begrippen en goede kennis hebben van het bedrijf c.q. bronsysteem
- Personen die het systeem hebben ontwikkeld, tezamen met gebruikersvertegenwoordiging moeten het testen begeleiden

Fout 8: Gebruik van 'zelf ontworpen' test data

- Maak gebruik van data sets uit productie voor alle testen, zeker bij systeem- en gebruikers acceptatie testen
- Als 'echte' data niet beschikbaar is, dan moet de implementatie van het data warehouse worden uitgesteld. Het testen met productie data is zo belangrijk

Traditioneel testen versus BI testen

Waterval testen versus BI testen

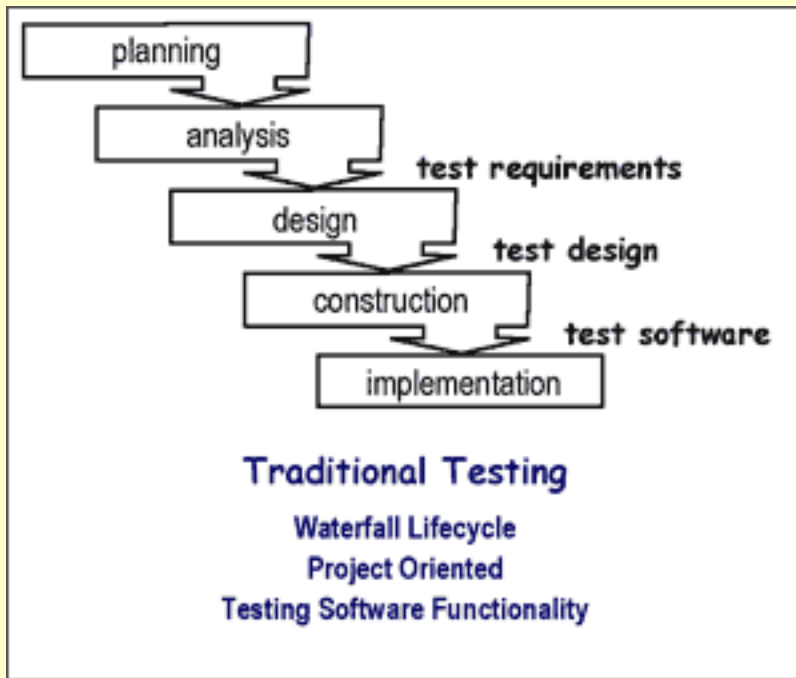


Figure 1: The Waterfall Lifecycle

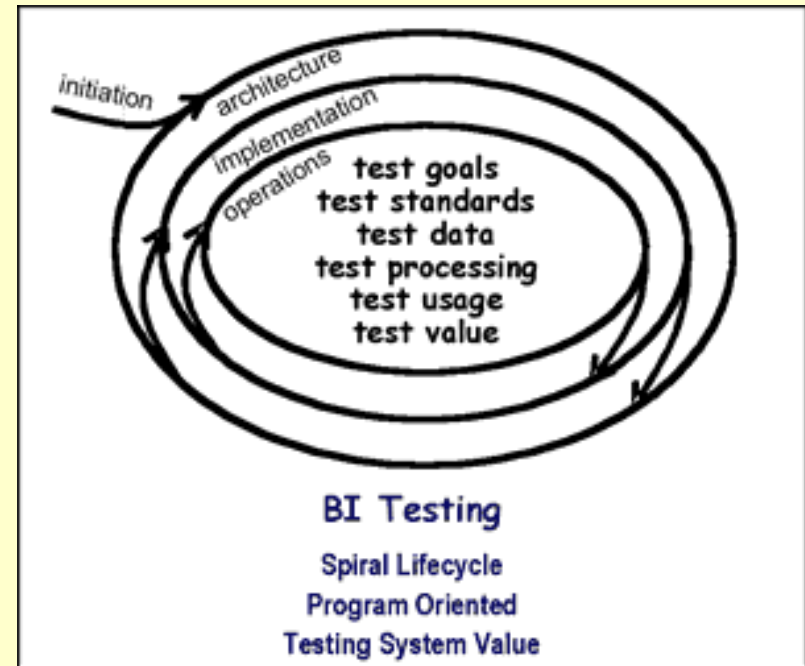


Figure 2: The BI Lifecycle



TDWI FlashPoint: Testing Techniques for BI Systems, 15 dec 2004

Testen (1)

- Er zijn diverse verschillen die bijdragen tot de uitdaging van DWH/BI testen
 - Business requirements
 - Analyse van software ontwikkeling
 - BI requirements
 - Project versus programma
 - Een typisch software ontwikkelingsproject staat op zichzelf
 - BI is een voortdurend programma zonder vast eind

Testen (2)

Nadruk van het testen

-  Software testen heeft een enkelvoudig doel
-  BI testen legt de nadruk op de bruikbaarheid van de informatie die geleverd wordt

Objecten van het testen

-  Software testen is hoofdzakelijk testen van programma code
-  BI testen wordt gedirigeerd door data en informatie

Testen (3)

- **Proces van het testen**
 - Software testen → Alle code kan getest worden (unit- en systeemtest) voordat de software gereleased wordt
 - Vergelijkbare volledigheid van BI testen voordat wordt gereleased is onpraktisch

Deskundig inzicht over DWH testen

(http://www.dmreview.com/article_sub.cfm?articleId=5701)

- Testen van het ETL proces is de primaire doelstelling. Alle rijen die je wilde laden moeten ook geladen zijn
- De aandachtsgebieden die getest moeten worden zijn de ETL processen, SQL queries, presentatie tool, stored procedures of queries die aggregaat tabellen maken
- Een belangrijk punt van het testen van een data warehouse is dat je de data kent en ook de antwoorden die je verwacht
- Testen zal nooit de garantie geven dat er nooit meer een fout wordt gevonden

Typisch test proces voor Data Warehouse

- Unit test
- System test
- Acceptatie test

Unit test

- Checken van de ETL procedures/mappings/jobs en de ontwikkelde rapportages
- Beschouw dit als een white-box test

Systeme test

- Unit test moet zijn afgerond en vormt de basis voor de systeem test
- Het systeem wordt getest met volledige functionaliteit en de verwachting is dat het functioneert zoals het in productie moet functioneren
- Op het eind van deze fase moet het systeem acceptabel zijn voor de klant
 - In termen van ETL: proces integriteit en bedrijfsfunctionaliteit
- Om deze activiteit uit te kunnen voeren moet het testteam kennis hebben van de functionaliteit van het bedrijf
- Valideren dat de juiste data is geladen vanuit het bronsysteem in het data warehouse

Ook white-box test

Acceptatie test

- **Systeem test moet zijn afgerond en vormt de basis voor de acceptatie test**
- **Focus bij acceptatie test zijn de rapportages**

Black-box test

Testen van validiteit van data in het rapport met behulp van bron systeem

- De integriteit van data die wordt getoond in rapportages is cruciaal voor acceptatie
- Het testteam moet het ETL proces beschouwen als een 'black-box'

Performance testen

■ Back-end

- Tijd voor het verversen van standaard rapporten
- Tijd voor het verversen van complexe rapporten
- Er moet voldoende data aanwezig zijn in het data warehouse om representatief te zijn
- Laden van data voor een enkele dag als ook batch windows moeten worden getest (gereed binnen laad 'window')

■ Front-end

- Hoe lang duurt het om met een aantal gebruikers eenzelfde rapport op te vragen

Meer weten?

Centennium BI expertisehuis
Lange Voorhout 43
2514 EC 's-Gravenhage
Telefoon 070 31 20 370
Fax 070 31 20 371
URL www.centennium.nl



- **Anita van Berghenegouwen**
business intelligence consultant
a.vanberghenegouwen@centennium.nl

- **Eric Hoogenboom**
manager sales & delivery
e.hoogenboom@centennium.nl

- **Erik Fransen**
senior business consultant / manager business development
e.fransen@centennium.nl