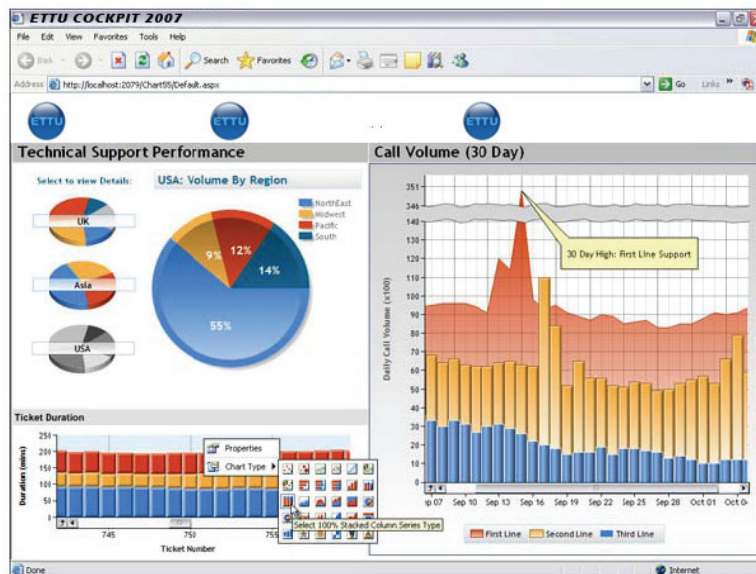


Afstudeerverslag – Business Intelligence



Ronald Kraijesteijn, juni 2007



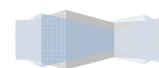
Afstudeerinformatie

Naam: Ronald Kraijesteijn
Studentnummer: 20040698
Email: kraaitje@tiscali.nl

Afstudeeropdracht: Business Intelligence
Afstudeerperiode: 2007-feb – 2007-juni
Verslagdatum: 8 juni 2007, versie 1.0

Opdrachtgever: ETTU
Mentor en opdrachtgever: V. Heijckmann
Adres: Wilhelmina van Pruisenlaan 139
2807 KH Gouda
info@ettu.nl
Tel: 0182-686000

Opleiding: HBO Bedrijfskundige Informatica
Instelling: Academie voor ICT & Media, Zoetermeer
Examinatoren: A. Nederend
A.J. Warnink



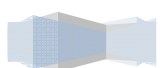
REFERAAT

Dit procesverslag is geschreven in het kader van het afstuderen van de opleiding Bedrijfskundige Informatica aan de Academie voor ICT & Media te Zoetermeer. Het behandelt de werkzaamheden die zijn verricht bij de firma ETTU, gevestigd in Gouda. Hier heb ik mijn afstudeeropdracht voltooid in de periode februari – juni 2007.

De opdracht betrof het opzetten van een managementdashboard waarin diverse Key Performance Indicators (KPI's) en rapportages waren af te lezen om inzicht te geven in de situatie voor het komende jaar.

Descriptoren:

- Business Intelligence;
- ETL;
- Sterschema;
- Kimball;
- Managementinformatie;
- Succesfactor;
- Datawarehouse;
- Managementdashboard;
- Key Performance Indicator/Prestatie-indicator/KPI;
- Microsoft Sharepoint 2007;
- Analysis Services;
- Integrated Services;
- Reporting Services;
- Excel services;
- Microsoft SQL Server 2005.



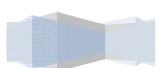
VOORWOORD

Tijdens de opleiding Bedrijfskundige Informatica wordt in het laatste half jaar door de student afgestudeerd. In deze periode wordt er van de student verwacht dat hij de in voorgaande jaren opgedane kennis in de praktijk kan toepassen en uitbreiden. Hiervoor dient de student een bedrijf te vinden dat de mogelijkheden hiertoe biedt.

Voor het afstuderen heb ik een bedrijf uitgekozen waar ik mijn opgedane kennis optimaal kon benutten en waar er ruimte was voor een eigen inbreng. ETTU is een jonge en innovatieve organisatie die mij deze kans bood. Dit rapport is het resultaat van het afstudeerproject “Business Intelligence” bij de firma ETTU. Dit rapport bevat de beschrijving van de ontwikkeling van een zogeheten managementdashboard met bijhorende technische ontwikkelingen. Dit geheel van applicaties en technologieën voor het verzamelen, opslaan en analyseren van gegevens waarop bedrijfsstrategieën worden gebaseerd heet Business Intelligence.

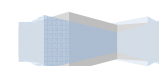
Ik zou graag alle medewerkers van ETTU willen bedanken voor de gezellige en leerzame afstudeerperiode. Verder wil ik mijn begeleiders Arno Nederend en Mieke Warnink bedanken voor hun begeleiding tijdens het afstuderen. Ook bedank ik Peter Ritman voor zijn begeleiding tijdens mijn studieloopbaan. Tot slot bedankt ik de Academie voor ICT & Media in Zoetermeer voor de fijne tijd op school.

Ronald Kraijesteijn, juni 2007

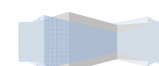


INHOUDSOPGAVE

1.0	INLEIDING	7
2.0	ACHTERGROND INFORMATIE	8
2.1	Inleiding	8
2.2	Het bedrijf ETTU	8
2.3	Succesfactoren binnen ETTU	9
2.4	Huidige voorziening managementinformatie	9
2.5	Gewenste situatie managementdashboard	10
3.0	PROJECTPLAN ONTWIKKELEN MANAGEMENTDASHBOARD	12
3.1	Inleiding	12
3.2	Opdracht en doelstelling	12
3.3	Methode en activiteiten	13
3.4	Op te leveren producten	15
3.5	Globale planning.....	15
3.6	Projectrisico's	16
4.0	FASE 1: JUSTIFICATION	18
4.1	Inleiding	18
4.2	Analyseren business van ETTU	18
4.3	Doelstellingen Business Intelligence in kaart brengen.....	19
5.0	FASE 2: PLANNING.....	20
5.1	Inleiding	20
5.2	Betrokkenheid organisatie	20
5.3	Afbakening en milestones	21
5.4	Infrastructuur	21
5.5	Personeel en ervaring.....	22
6.0	FASE 3: BUSINESS ANALYSIS	23
6.1	Inleiding	23
6.2	Project Requirements opstellen	23
6.3	Data-analyse	29
6.4	Application prototyping.....	36
6.5	Meta-data repository design	37
7.0	FASE 4: DESIGN DATAWAREHOUSE, ETL-SCHEMA'S, DASHBOARD	38
7.1	Inleiding	38
7.2	Opstellen bronnenmodel	38



7.3	Opstellen gegevensmodel	39
7.4	Opstellen feitmodel	41
7.5	Opstellen sterschema	43
7.6	Vorbereidende activiteiten ETL-schema's.....	45
7.7	ETL-schema budgetsysteem	49
7.8	ETL-schema urensysteem	52
7.9	ETL-schema CRM-systeem.....	52
8.0	FASE 5: CONTRUCTION	54
8.1	Inleiding	54
8.2	Literatuurstudie en toolverkenning SQL Server 2005	54
8.3	Toolverkenning Microsoft Sharepoint 2007.....	57
8.4	Constructie ETL-packages	57
8.5	Constructie datawarehouse	58
8.6	Constructie Dashboard in Microsoft Sharepoint 2007	60
9.0	FASE 6: DEPLOYMENT	63
9.1	Inleiding	63
9.2	Testen	63
9.3	Implementatie	64
10.0	EVALUATIE.....	65
10.1	Inleiding	65
10.2	Evaluatie eindproducten	65
10.3	Evaluatie proces.....	66
10.4	Evaluatie planning	67
BIJLAGE A: OPDRACHTOMSCHRIJVING		69
BIJLAGE B: PLAN VAN AANPAK		72
BIJLAGE C: BUSINESS REQUIREMENTS REPORT		81
BIJLAGE D: BUSINESS DESIGN REPORT		92
BIJLAGE E: IMPLEMENTATION REPORT		112
BIJLAGE F: SCREENSHOTS DATAWAREHOUSE		126
BIJLAGE G: SCREENSHOTS MANAGEMENTDASHBOARD.....		127
BIJLAGE H: SCREENSHOTS ETL-PACKAGES		128
BIJLAGE I: FIGURENLIJST		129



1.0 Inleiding

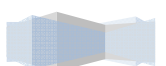
Dit verslag bevat de proces- en productverantwoording voor de door mij uitgevoerde afstudeeropdracht 'Business Intelligence' bij de firma ETTU.

Door middel van dit verslag zal ik proberen de door mij uitgevoerde werkzaamheden te verantwoorden. Daarbij heb ik rekening gehouden met de doelgroep van dit verslag; de geëncmitteerden van de Haagse Hogeschool. Ik heb bij het schrijven van dit verslag geprobeerd mij te verplaatsen in de lezer. Dat betekent dat ik er rekening mee heb gehouden dat de lezer niets weet van zowel de opdracht, het bedrijf als het probleem. Ik ben er wel vanuit gegaan dat de lezers enige kennis bezitten van het betreffende onderwerp/vakgebied Business Intelligence.

Het verslag is opgedeeld in een aantal hoofdstukken. Na deze inleiding volgt het hoofdstuk 'achtergrond' waarin meer informatie te vinden is over het bedrijf ETTU, de succesfactoren en de huidige voorziening in managementinformatie. Daarna volgt een hoofdstuk over de projectaanpak waarin uiteen wordt gezet hoe de aanpak van het afstuderen was.

De activiteiten van het vervullen van het project zijn onderverdeeld in een zestal fasen. Deze fasen vormen de kern van het afstudeerproject en worden behandeld worden in de daaropvolgende hoofdstukken.

Het laatste hoofdstuk behandelt mijn evaluatie op de eindproducten, het proces en de planning.



2.0 Achtergrond informatie

2.1 Inleiding

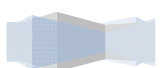
Deze paragraaf behandelt enige achtergrondinformatie betreffende de firma ETTU en de behoefte tot het starten van dit project. Na het lezen van dit hoofdstuk moet er meer duidelijkheid zijn ontstaan over wat voor type organisatie ETTU is, wat de succesfactoren van ETTU zijn en hoe men op dit moment voorziet in managementinformatie. Daardoor wordt ook duidelijk waarom dit project is gestart en wat de gewenste situatie is na afloop van dit project.

2.2 Het bedrijf ETTU

ETTU is een jonge en innovatieve organisatie die in 2006 is opgericht en waar op dit moment ongeveer 20 mensen werkzaam zijn. ETTU is innovatief omdat men inspeelt op de nieuwste marktontwikkelingen en trends. ETTU is ook een snelgroeiende organisatie, dat blijkt wel uit het feit dat het bedrijf in een jaar tijd is gegroeid met 15 mensen. Het management van ETTU bestaat uit drie leden die eigenaar zijn van de aandelen van het bedrijf. Het bedrijf is opgesplitst in drie businessunits; Applications, Configurations en Detachering. De laatste is erg in ontwikkeling. Er is kort geleden een personeelslid bijgekomen die mensen gaat rekruten voor deze businessunit. De eerste businessunit ontwikkelt voornamelijk mobiele toepassingen. Hierbij valt te denken aan systemen die gebruikt worden voor het maken van digitale werkbonden voor monteurs.

De businessunit Configurations ontwikkelt Sharepoint oplossingen en levert consult op dit gebied. Deze oplossingen worden gebruikt om documenten binnen organisaties beheersbaar te maken. Ook zijn deze oplossingen erg flexibel en kunnen volledig worden aangepast aan de eisen en wensen van de klant. Wat er nog niet is kan gemaakt worden door de kennis van .NET (dot net). ETTU probeert zichzelf te onderscheiden van andere partijen door productleader te zijn op deze twee gebieden. Verder maakt zij gebruik van de tien gouden regels die zowel intern als extern gelden, de ADA-factoren (Anders dan Anderen).

De motor van ETTU draait op de verkoop van declarabele uren en projecten. ETTU werkt met een budget en een forecast. Het budget wordt aan het begin van ieder kalenderjaar vastgesteld. Hierin wordt budget gereserveerd voor alle uitgaven per businessunit. De forecast lijkt sterk op het budget, alleen worden in de forecast de werkelijke cijfers vastgelegd en worden voorspellingen gedaan richting het huidige jaar. In de forecast wordt bijvoorbeeld vastgelegd wat de omzet aan het einde van een maand is geweest. Ook wordt hierin vastgelegd dat er in een bepaalde maand een medewerker wordt aangenomen. Het verschil tussen de forecast en het budget is dat het budget vastligt en niet kan worden gewijzigd. De forecast wordt continu bijgesteld. Het streven van ETTU is om binnen budget te blijven als er gekeken wordt naar de kosten. Daarnaast streeft men er uiteraard naar om meer winst te maken dan is vastgelegd in het budget.



2.3 Succesfactoren binnen ETTU

Kritische succesfactoren geven aan op welke gebieden ETTU moet excelleren om de strategische doelen te realiseren. Deze factoren zijn afgeleid uit het missiestatement welke staat omschreven in het businessplan van ETTU. De succesfactoren die voor ETTU gelden zijn:

1. Betrouwbaar zijn;
2. Winstgevend zijn;
3. Klantgericht zijn;
4. Uitstekende reputatie hebben;
5. Leveren van consultancy en development;
6. Hoog opgeleid personeel hebben;
7. Personeel zich thuis laten voelen;
8. Open, transparant en innovatief zijn;
9. Garant staan voor de kwaliteit van opgeleverde oplossingen;
10. Investeren van de klant waarborgen.

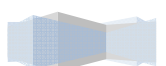
Dit project zal zich bezighouden met de tweede succesfactor, winstgevend zijn. In een vroeg stadium van dit project is in overleg met de opdrachtgever besloten om deze succesfactor te gaan meten. De andere factoren zijn zeker niet minder belangrijk maar er is bewust gekozen om juist de tweede te meten omdat de motor van ETTU draait op winst. Wanneer dit project een succes blijkt te zijn kan ETTU besluiten om de overige succesfactoren te gaan meten.

2.4 Huidige voorziening managementinformatie

Er is niet zonder reden besloten om dit project te starten. Managementinformatie is belangrijk om de juiste beslissingen te kunnen nemen, voornamelijk op tactisch en strategisch niveau. Het management van ETTU is voornamelijk geïnteresseerd in informatie die gaat over geld, uren en projecten. Enkele voorbeelden van informatie waarin het management van ETTU is geïnteresseerd: declarabiliteit van medewerkers, voorspelde winstgegevens en de status van een project. Juist die zaken waar het succes van ETTU door gevormd wordt zou men graag beschikbaar hebben in de vorm van een rapportage of een metertje. Voor ETTU was er aanleiding voor het starten van dit project omdat deze informatie op dit moment niet voorhanden is of omdat deze informatie niet beschikbaar is.

Momenteel kan men in sommige informatie voorzien, alleen is het niet voorhanden. Men moet in de huidige situatie gegevens die in verschillende systemen zijn vastgelegd met elkaar vergelijken om een conclusie te kunnen trekken. Wil men weten hoeveel budget er is opgebruikt voor een project dan zal men alle geschreven uren voor het betreffende project bij elkaar op moeten tellen en moeten vermenigvuldigen met diverse uurlonen van medewerkers. Daarna moet men gaan kijken hoeveel budget er voor dit project is en zal men dit handmatig moeten vergelijken. Uit dit voorbeeld blijkt dat de gewenste informatie niet voorhanden is.

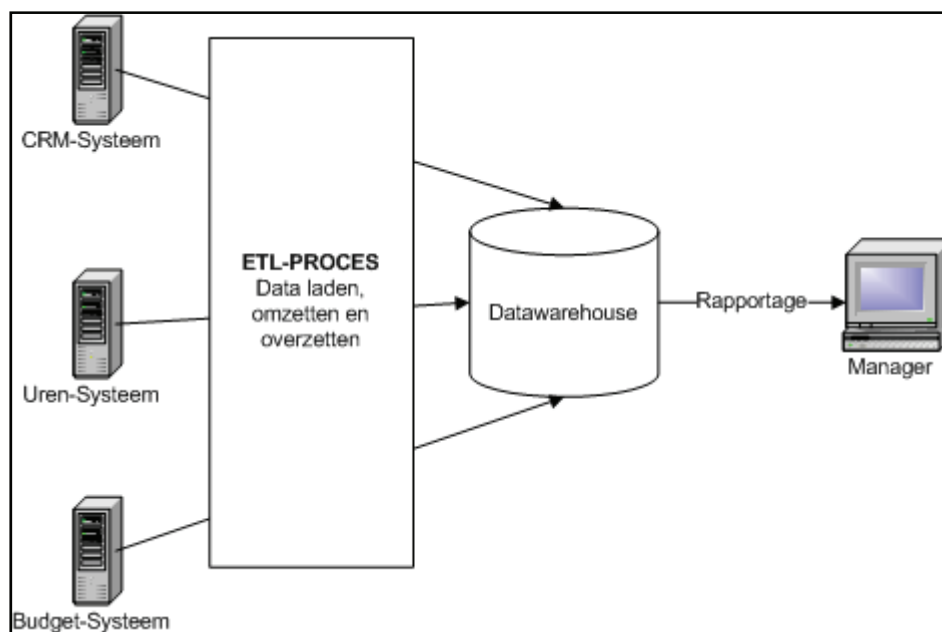
Sommige informatie kan op dit moment niet worden verkregen. ETTU werkt met een CRM-systeem waarin bestellingen worden bijgehouden en waar ook de opportuniteiten in worden bijgehouden. Men heeft op dit moment geen zicht op de waarde van de bestellingen die in een bepaalde maand in deze portefeuille aanwezig is. Ook het overzicht van opportuniteiten voor een bepaalde maand is niet aanwezig.



2.5 Gewenste situatie managementdashboard

Zoals uit bovenstaand voorbeeld is af te leiden is het soms lastig om de juiste gegevens naar voren te toveren. Voor ETTU zou het interessant zijn om alle benodigde gegevens die op dit moment lastig kunnen worden verkregen te verzamelen in een database en vervolgens de interessante data te presenteren in een dashboard. De gewenste situatie na afloop van dit project is daarom een managementdashboard met metertjes, lampjes en cijfers waarmee de managers in een oogopslag kunnen zien hoe ETTU ervoor staat. In dit dashboard zou ETTU de volgende informatie willen aflezen:

- Financiële informatie;
- Projecteninformatie;
- Informatie over bestellingen.



Figuur 1 Business intelligence bij ETTU

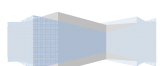
Financiële gegevens

Met betrekking tot financiële stuurinformatie wil ETTU met behulp van een stoplicht of lijngrafiek kunnen zien wat de verhouding is tussen forecast- en budgetgegevens. Een hoger forecast-gegeven kan zowel positief als negatief zijn. Als voorbeeld worden de directe kosten genomen. Als blijkt dat de voorspelde directe kosten (forecast) voor de volgende maand opeens omhoog schieten als gevolg van een wijziging in het budgetsysteem dan valt dit direct op doordat het lampje in het dashboard rood gaat branden. Daarnaast is de verhouding met het budget direct te zien met behulp van een lijngrafiek die boven de forecast uitstijgt.

Veranderingen in de forecast hebben gevolgen voor de bedrijfswinst, de EBIT-gegevens (Earnings before interest and taxes). Ook voor deze gegevens is een budget en een forecast door ETTU opgesteld. Wanneer de forecast daalt onder de gebudgetteerde winst dan zal het stoplicht ook op rood springen.

Projecteninformatie

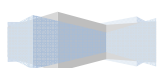
ETTU wil graag inzichtelijk hebben hoeveel budget er voor een project gebruikt is. Dit gebruikte budget ontstaat vanaf het moment dat de eerste uren voor een project zijn geschreven t/m de laatst geschreven uren. Dit kan uitgerekend worden door het aantal gewerkte uren te vermenigvuldigen



met het uurloon. Zodra er meer budget gebruikt is dan er voor het project is berekend dan moet het lampje op rood springen.

Informatie over orders

ETTU zou graag inzichtelijk hebben wat het bedrag is aan bestellingen dat in een bepaalde maand in portefeuille aanwezig is. Verder wil men ook graag zien wat het bedrag is aan opportuniteiten dat in diezelfde maand aanwezig is. Deze twee moeten worden afgezet tegen de maximale omzet van de eigen medewerkers. Deze gegevens wil men vervolgens zien voor het alle maanden van het huidige jaar. Men wil dus in januari van 2007 willen zien welk bedrag er voor de maand augustus 2007 in portefeuille aanwezig is. Men heeft hierdoor beter zicht op de prospectieve situatie van het komende jaar. Blijkt dat er in juni verschrikkelijk veel orders gaan vallen maar er onvoldoende mankracht is (omzet medewerkers) kan men besluiten een extra FTE (Full Time Employee) aan te nemen.



3.0 Projectplan ontwikkelen managementdashboard

3.1 Inleiding

Het plan van aanpak was het eerste document dat voor dit project is opgesteld. In dit plan van aanpak zijn diverse belangrijke aspecten van het project behandeld. Het bevat een omschrijving van de opdracht met afbakening en risico's. Verder zijn alle fasen die voor dit project doorlopen moesten worden beschreven met bijhorende milestones. Om een indruk te geven van de opbouw van dit plan is de inhoudsopgave opgenomen.

1.0	INLEIDING	5
2.0	PROJECTDEFINITIE	6
§2.1	Doelstelling	6
§2.2	Aanpak en fasering	6
§2.3	Fasen en activiteiten	7
§2.4	Milestones	8
§2.5	Afbakening	9
§2.6	Methoden en technieken	10
3.0	BUSINESS CASE	11
§3.1	Aanleiding	11
§3.2	Risico's	12
§3.3	Succesfactoren	12
4.0	PLANNING	13
5.0	PROJECT BEHEERSING	15
§5.1	Voortgangsrapportages	15
§5.2	Voorwaarden aan derden	15
§5.3	Communicatie en vastlegging	15

Figuur 2 Plan van aanpak

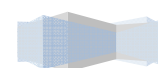
3.2 Opdracht en doelstelling

De opdracht luidde als volgt: ontwikkel een managementdashboard waarin diverse KPI's worden weergegeven en integreer dit in de bestaande Microsoft Sharepoint omgeving van ETTU.

Voordat het afstuderen is begonnen heb ik aan de opdrachtgever een document gevraagd waarin hij zijn ideeën over het project kwijt kon. Zo werd het voor mij concreet waarom het project uitgevoerd moest worden en wat hij van het project verwachtte. Ik heb aan hem gevraagd of hij het volgende kon documenteren:

- Aanleiding;
- Doelstelling;
- Milestones;
- Concrete werkzaamheden.

Het doel van het project was duidelijk. ETTU had behoefte aan een managementdashboard binnen hun bestaande intranet (Sharepoint 2007) omgeving waarin diverse prestatie-indicatoren waren af te lezen. Daarbij valt te denken aan projectvorderingen, budgetten en uurtarieven. Deze informatie was



zoals eerder is vermeld niet beschikbaar of niet voorhanden. Door deze informatie beschikbaar te maken zou het management in staat zijn betere beslissingen te nemen.

De doelstelling van het project is in vroeg stadium vastgelegd in het plan van aanpak. Het doel van het plan van aanpak was het vaststellen van onder meer:

- De definitieve (afgebakende) opdracht;
- De gehanteerde methodiek;
- De planning.

3.3 Methode en activiteiten

Om het project beter te kunnen beheersen is er gebruik gemaakt van een methode. De keuze is gevallen op de methode uit het boek: Business Intelligence Roadmap, The Complete Lifecycle. De reden voor de keuze van deze methode heeft meerdere argumenten. Het boek maakt gebruik van best practices. Dat betekent dat het vertelt hoe je het zou kunnen doen, niet hoe je het moet doen. Er zijn meerdere manieren om een B.I. project te voltooien en dit boek beschrijft met welke punten je rekening moet houden. Dit is ook direct een zwakke kant van het boek. Het is erg uitgebreid en daardoor is het soms onduidelijk welke stappen er in ieder geval gedaan moet worden en welke stappen optioneel zijn.

De methode behandelt als eerste de complexiteit van B.I. (Business Intelligence) projecten. Vervolgens presenteert de methode een stappenplan voor het gehele B.I. project welke is verdeeld in zes fasen. Ieder hoofdstuk is voorzien van afbeeldingen waarin af te lezen is in welke fase de lezer zich bevindt. Bij iedere fase is een checklist te vinden waarop de lezer moet letten als hij zich in die fase bevindt en is er te lezen welk risico er is wanneer die fase niet wordt doorlopen.

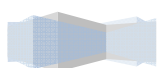
Andere zaken die mij aanspraken bij het bekijken van de methode die in dit boek stond beschreven:

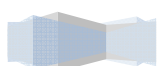
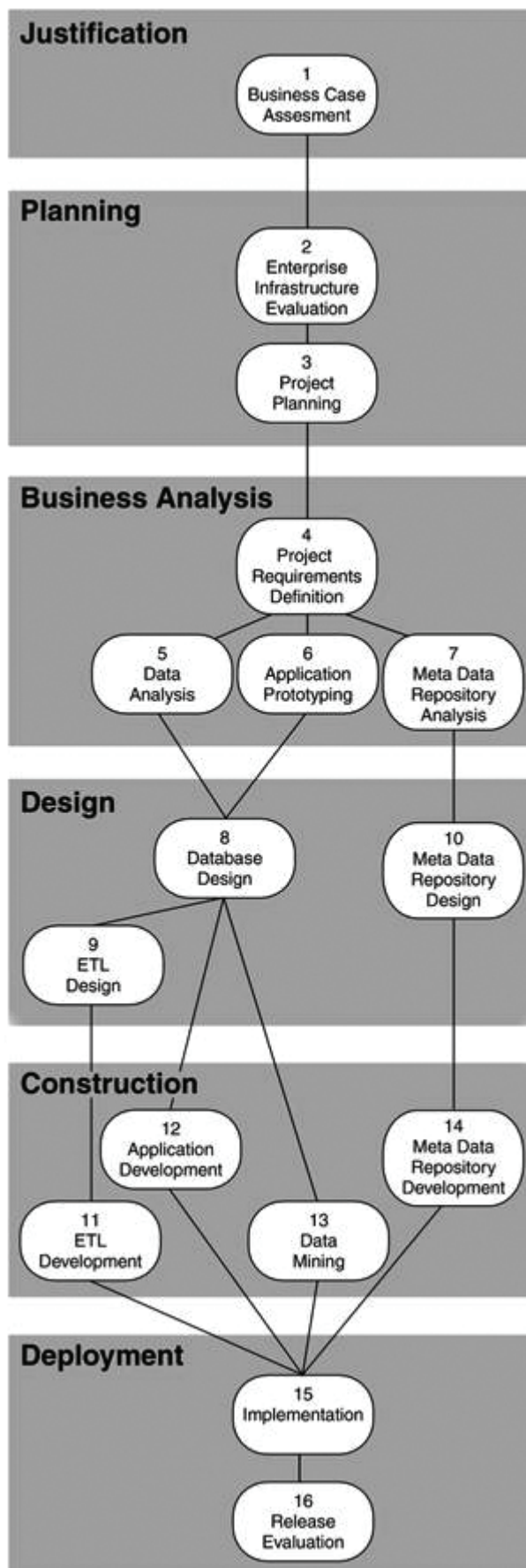
- Matrix met betrokken personen per fase;
- “Entry, Exit, Deliverables” matrix waarin per fase is beschreven hoe je de fase in- en uitgaat en welke producten er in die fase worden opgeleverd.

Niet alle zaken die deze methode behandelt heb ik gebruikt in mijn project. De reden daarvan is de omvang van het project en het tijdsbestek. Er is niet genoeg tijd om alle producten te maken die in de fasen worden beschreven en niet alle producten zijn van toepassing op dit project. De keuze om onderdelen van een fase weg te laten zal ik toelichten bij het behandelen van de desbetreffende fase.

Voor de ontwikkeling van het datawarehouse zijn er andere stappen gebruikt die een goede houvast hebben geboden. De reden om deze stappen te volgen was het feit dat er in de methode geen duidelijke stappen waren beschreven voor de ontwikkeling van een datawarehouse. Voor dit project zal de kennis die benodigd is van de ontwikkeltools worden opgedaan door het volgen van Microsoft tutorials.

De fasen die voor dit project zijn gebruikt zijn te vinden in onderstaand figuur op de volgende pagina.





3.4 Op te leveren producten

Dit project heeft een aantal producten opgeleverd. Het was duidelijk dat de opdrachtgever een managementdashboard wilde dat de juiste cijfers moest presenteren. Omdat gebruik werd gemaakt van bovenstaande methode waren er meer milestones die opgeleverd moesten worden. Het eerste document dat opgeleverd zou worden was het plan van aanpak. Deze milestone stond niet in de methode vermeld maar geldt voor ieder project. In dit plan van aanpak stonden alle belangrijke gegevens over het project en vormde als het ware een contract tussen mij en de opdrachtgever. Er zijn afspraken in te vinden zoals welke producten opgeleverd moesten worden, de afbakening en de planning.

Daarna moest er onderzocht worden aan welke informatie het management en de business behoefte hadden. Deze zogeheten businessneeds zijn vastgelegd in een Business Requirements Report. Dit is een milestone van fase 3; business analysis. Dit document vormt het uitgangspunt voor het te ontwerpen systeem. In dit document is de informatiebehoefte terug te vinden, een opzet van de gewenste rapportages, uitwerking van de KPI's (Key Performance Infactors) en de datakwaliteit.

De volgende milestone laat alle ontwerpen zien van het te bouwen systeem. Dit document is de output van fase 4; design. In dit rapport is naast het datawarehouseontwerp ook een ontwerp te vinden van de diverse ETL-processen.

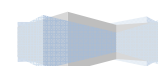
De laatste milestone was een werkende oplossing, gebouwd in Business Intelligence Studio in combinatie met SQL Server 2005 en Sharepoint 2007.

Een extra milestone die uiteindelijk is opgeleverd is het implementation report. Dit rapport bevat informatie over hoe de oplossing geïmplementeerd moet worden.

3.5 Globale planning

Hieronder is de globale planning te vinden. De planning is opgedeeld in de zes fasen die doorlopen zijn en die afgeleid zijn uit eerder genoemde aanpak. Achter iedere fase zijn de milestones te vinden die de output zijn van de betreffende fase. Een gedetailleerde planning voor dit project is opgenomen in het plan van aanpak in de bijlage. De benodigde tijd per fase is ontstaan door een ruwe schatting.

Fase	Output	Geplande startdatum	Geplande einddatum
1.0 Justification	Business Case	5-2-2007	9-2-2007
2.0 Planning	Plan van aanpak	8-2-2007	14-2-2007
3.0 Business analyses	Business Requirements report	8-2-2007	2-3-2007
4.0 Design	Business Design Report	5-3-2007	23-3-2007
5.0 Construction	Datawarehouse, ETL packages, Dashboard	23-3-2007	1-6-2007
6.0 Deployment	Implementation Report	4-6-2007	8-6-2007



Dit project is gedeeltelijk lineair uitgevoerd en gedeeltelijk iteratief. Dit heeft een nadere uitleg nodig. De eerste drie stappen zijn lineair uitgevoerd. Stap vier kon ook niet worden gestart zonder dat de eerste drie stappen waren doorlopen. Daarna is er telkens een stukje systeem gebouwd met het idee: beter een klein stukje systeem helemaal werkend dan op het laatst alles een beetje werkend. De stukjes systeem zijn gebaseerd op de drie bronsystemen. Zo is ervoor gekozen eerst het budgetstelsel uit te werken (prioriteit 1), vervolgens de orderportefeuille en tot slot de project-budgetverbruik. Zodra bleek dat een stukje systeem werkte is het volgende stukje ontworpen en gebouwd. Tot slot is alles samengevoegd in een werkend systeem. Voordeel hiervan was ook dat fouten in het systeem eenvoudiger konden worden gevonden dan wanneer het systeem erg groot zou zijn.

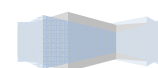
Zoals af te leiden is uit de planning heb ik de meeste tijd gereserveerd voor fase 5. Er is bewust extra tijd ingelast omdat er nog geen kennis aanwezig was van de software. Uit ervaring is gebleken dat het leren omgaan met nieuwe applicaties veel tijd kost. Daarnaast was er binnen ETTU geen kennis van Business Intelligence. Dit betekende dat ik op mijzelf aangewezen was.

3.6 Projectrisico's

In vroeg stadium is er nagedacht over de risico's die een rol konden spelen binnen dit project. Buiten de standaard risico's die voor ieder project gelden zoals ziekte en dataverlies, was ik ervan overtuigd dat er meer risico's aanwezig waren. Ik heb rondgekeken op internet naar risico's die de meeste Business Intelligence projecten doen mislukken. Het bleek namelijk dat er voor B.I. projecten nogal wat risico's aanwezig waren. Business Intelligenceprojecten zijn projecten waar vaak een grote investering voor nodig is. Omdat dit niet van toepassing was voor dit project maar omdat ik geïnteresseerd was in deze materie ben ik gaan zoeken op internet naar projectrisico's. Daar kwam ik een aantal herkenbare zaken tegen die ik van toepassing vond op mijn project. De belangrijkste risico's op een rijtje:

- Planning: het bleek uit de praktijk dat B.I. projecten vaak uitlopen en daardoor niet op tijd opgeleverd kunnen worden. Omdat er slechts 15 weken beschikbaar zouden zijn voor de uitvoer van het project zou dit een potentieel risico kunnen vormen. Een goede maatregel: **prioriteiten stellen**: eerst de zaken behandelen die het belangrijkste zijn en dus een prioriteitenlijst hanteren. Dit is gedaan door een top vijf te hanteren van KPI's.
- Datakwaliteit: bij B.I. draait het om beslissingen nemen aan de hand van voorgeschiedte rapportages. De tijd dat bedrijven in zekere mate afhankelijk zijn van IT is voorbij. Men is tegenwoordig volledig afhankelijk van IT. Het management moet dus blindelings kunnen vertrouwen op de techniek. Wanneer de rapportages of metertjes niet kloppen kan dit ertoe leiden dat er verkeerde tactische of strategische beslissingen worden genomen, hetgeen funest kan zijn. Ik vond het belangrijk om dit als risico mee te nemen zodat ik er tijdens de uitvoering van het project extra alert op moest zijn dat de data zou kloppen. Een goede maatregel: medewerkers bewust maken van goede datakwaliteit en zorgen dat het systeem na afloop getest wordt op juiste gegevensuitvoer of KPI.

In onderstaande tabel zijn de risico's te vinden. Er zijn ook maatregelen in opgenomen om te zorgen dat om het risico te voorkomen.

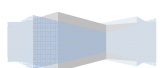


Tabel 1: Risicotabel			
Risico	Omschrijving	Risico	Maatregel
Technologie	Weinig ervaring met de technologie	▲	Zelfstudie, goede inzet
Planning	Te krap plannen, project niet op tijd klaar	◆	Planning in deelfasen, werk beter inschatten, beperken tot de primaire processen
Integratie	Integreren binnen SharePoint geeft problemen	▲	Hulp inschakelen van expert op dit gebied
Support	Weinig kennis binnen de organisatie op het gebied van business Intelligence	▲	Goede zelfstandigheid van de uitvoerder, support op school vragen
Communicatie over en weer	Te weinig interactie kan ervoor zorgen dat de wens van de opdrachtgever niet goed is afgestemd met het eindresultaat	▲	Wekelijks vergadermoment inbouwen in de planning
Datakwaliteit	Kwaliteit data slecht of niet integer, daardoor kloppen de rapportages niet	◆	Bewust maken van het belang van goede datakwaliteit, controleren op kwaliteit

▲ Gemiddeld risico

◆ Hoog risico

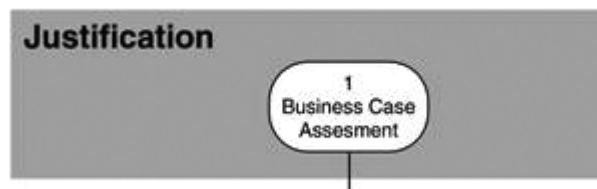
Figuur 3 Projectrisico's



4.0 Fase 1: Justification

4.1 Inleiding

De eerste fase die de methode beschrijft is de Justification fase, in het Nederlands “motivatie-fase”. In deze fase ben ik gaan kijken waarom dit project zinvol zou kunnen zijn. In deze fase heb ik onder andere het plan van aanpak opgesteld waarin de business case te vinden is. Het was voor mijzelf erg nuttig om deze fase te doorlopen zodat ik zelf ook beter begreep waarom het project zinvol was. Ook is er in deze fase gekeken hoe de werkwijze van ETTU was. Ik moest zo snel mogelijk weten welke succesfactoren belangrijk waren en hoe men hierop kon sturen.



Figuur 4 Fase 1: justification

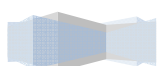
4.2 Analyseren business van ETTU

Aan het begin van mijn afstudeerperiode was het bedrijf ETTU voor mij een onbekend terrein. De probleemstelling was nog onduidelijk en het doel van het project was ook vaag te noemen. Men wilde iets met Business Intelligence gaan doen maar erg concreet was het idee nog niet. Het was noodzaak om wat meer over het bedrijf te weten te komen en de ideeën die men had. Om te kunnen bepalen waarom Business Intelligence een rol kon spelen binnen de organisatie was het belangrijk om een totaalplaatje voor mijzelf te creëren. Ik moest helder krijgen wat de huidige situatie was en wat de gewenste situatie was. Dit heb ik gedaan door te kijken hoe de werkwijze van ETTU was en hoe het bedrijf in elkaar stak. Zaken als bedrijfsstrategie en succesfactoren moesten duidelijk worden gemaakt. Het was niet nuttig om een Business Intelligence oplossing te ontwikkelen en te implementeren om te experimenteren met nieuwe technologieën maar het moest een zekere bijdrage leveren aan de business. De analyse moest antwoord geven op de volgende vragen:

- Wat is de missie van ETTU?
- Hoe werkt ETTU?
- Wat kan Business Intelligence bijdragen aan de strategie?

Een antwoord op deze vragen zou mij helpen te begrijpen wat de bijdrage kon zijn van dit project voor het bedrijf. Om een antwoord te krijgen op deze vragen heb ik aan de opdrachtgever te kennen gegeven wat meer over het bedrijf te weten te willen komen. Hij heeft aan mij het businessplan beschikbaar gesteld. In dit document was veel informatie te vinden over het bedrijf ETTU. Ook gaf dit document een antwoord op de vragen. Dit businessplan beschrijft onder andere de volgende zaken die interessant waren voor mijn project:

- Missie en strategie: nuttig voor het bepalen van succesfactoren;
- Key performance factors: nuttig om te kunnen bepalen waarop men wenst te sturen;
- Werkwijze per proces: een beeldvorming van de werkwijze van de bedrijfsprocessen.



Na het lezen van het businessplan was duidelijk wat de missie van ETTU was:

"ETTU is een betrouwbare, winstgevende en klantgerichte "product leader" met een uitstekende reputatie in het optimaliseren van de business van haar klanten middels consultancy en development in Microsoft Sharepoint, CRM, Bizstak en .NET technologie. Hoog opgeleide professionals voelen zich thuis bij de transparantie, openheid, bevoegdheid en innovatie van de organisatie en zullen zich betrokken voelen bij de klant en garant staan voor de kwaliteit van de opgeleverde oplossing, waardoor investeringen van de klant gewaarborgd zijn."

Deze missie zou later nodig zijn bij het opstellen van kritische succesfactoren.

In de tweede week is er een afspraak gepland met de proceseigenaar van het proces 'bestuur'. Hij was verantwoordelijk voor het systeem van budgetteren en forecasten. Hij heeft uitleg gegeven over de werking van het budgetsysteem en welke rapportages wellicht interessant konden zijn om in een dashboard weer te geven. Hij heeft mij ook uitgelegd waarom het nuttig zou zijn om dit in een dashboard weer te geven. Er is na afloop een vervolgspraak gemaakt om de informatiebehoefte voor zijn proces te bespreken.

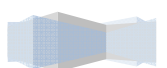
4.3 Doelstellingen Business Intelligence in kaart brengen

Om de tweede vraag te beantwoorden moest er dieper worden gegraven. Wat zijn de doelstelling van het toepassen van Business Intelligence binnen ETTU?

In het beginstadium heeft de opdrachtgever aangegeven waarom hij het nuttig vond om Business Intelligence te gebruiken. Op deze manier konden gegevens boven water worden gehaald die anders verborgen bleven of moeilijk te verkrijgen waren. Door het inzichtelijk maken van gegevens kon men sneller ingrijpen of bijsturen. Daarnaast was er nog een belangrijke reden om met managementinformatie en KPI's (Key Performance Indicators) te werken. Een van de directeurs heeft hiervoor een eigen bedrijf gehad dat een redelijke groei doormaakte. Een ander bedrijf dat in diezelfde periode gestart was groeide bijna 3x zo snel. De reden daarvoor was dat het management het bedrijf stuurde door gebruik te maken van KPI's.

Doelstellingen die behaald moest worden na voltooiën van dit project:

- Het sneller kunnen ingrijpen: directe- en indirecte kosten worden duidelijk weergegeven, kosten kunnen verlaagd worden indien dat nodig is.
- Het beter kunnen bijsturen: budgetverbranding per project wordt helder, stimuleren om effectiever met tijd om te gaan.
- Business Intelligence-functies in Microsoft Sharepoint 2007: nieuw in de 2007 versie van Sharepoint is de BI-functie. Ook biedt de nieuwe versie van Office BI-functies. ETTU wil graag kennis opdoen van deze nieuwe functies om dit mogelijk in de toekomst te kunnen verkopen aan klanten. Daarvoor wil ETTU eerst zelf experimenteren met Business Intelligence.

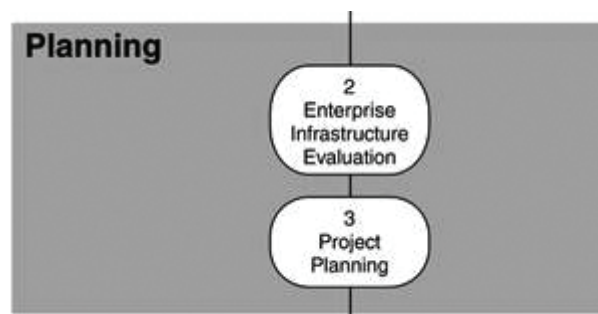


5.0 Fase 2: Planning

5.1 Inleiding

Bij het woord planning valt in eerste instantie te denken aan een projectplanning die tijdgebonden is. De planningsfase in dit project omvat echter meer dan alleen het plannen in tijd. De planningsfase behandelt namelijk niet alleen de planning van het project zelf maar behandelt ook andere onderwerpen waar over na moet worden gedacht:

- Betrokkenheid van de organisatie; communicatie met de opdrachtgever, beschikbaarheid van de opdrachtgever.
- Scope en milestones; hoe gedetailleerd zijn de eisen, wat wil de opdrachtgever graag zien en is dit haalbaar binnen de geplande periode?
- Kosten; kosten-baten analyse: overgeslagen, uitgaande dat dit project zijn toegevoegde waarde zal bewijzen na de implementatie.
- Infrastructuur; voldoet de infrastructuur, welke infrastructuur moet worden aangepast?
- Personeel en ervaring; wie zijn er betrokken bij het project, is er voldoende kennis?

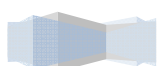


Figuur 5 Fase 2: planning

5.2 Betrokkenheid organisatie

De betrokkenheid van de organisatie achtte ik erg belangrijk voor het slagen van dit project. Het niet communiceren over het project zou kunnen leiden tot teleurstelling aan beide kanten, bijvoorbeeld omdat het resultaat niet overeen is gekomen met het verwachte resultaat. De betrokkenheid van de business is ook essentieel om tot de juiste modellen te komen. Ik kan wel van mening zijn dat een rapportage er op een bepaalde manier uit moet komen te zien maar ik ben uiteindelijk niet de persoon die met de rapportages zal gaan werken. Belangrijk was dus dat de opdrachtgever(s) betrokken waren bij het project. Betrokkenheid kon worden verkregen door overleg te hebben met de opdrachtgever over de voortgang en regelmatig het resultaat te tonen om aanpassingen op tijd door te kunnen voeren. Terugkoppeling richting opdrachtgever(s) was dus noodzakelijk, wellicht na iedere uitgevoerde fase van het project.

Een tweede belangrijk punt waar bij stil moest worden gestaan was de beschikbaarheid van de werknemers binnen de organisatie. ETTU is een bedrijf waarin gewerkt wordt met strakke plannings. Als er tijd nodig was bij een persoon moest dit (ruim) van tevoren in de agenda van deze persoon gepland worden. Dit was belangrijk om vooraf te weten. Verder moest er rekening mee worden gehouden dat ook na het inplannen van een afspraak de kans aanwezig was dat de afspraak



verzet zou worden. De voortgang van het project moest dus niet afhangen van een afspraak met een bepaalde persoon.

5.3 Afbakening en milestones

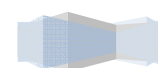
De afbakening van dit project is opgedeeld in een vijftal zaken: kwaliteit, budget, mensen en middelen, tijd en de scope. Bij de afbakening van het project is er rekening gehouden met twee dingen, de omvang en de diepgang. Het was de bedoeling om een evenwicht te vinden tussen beide zaken. Het project moest van voldoende omvang en diepgang zijn maar er moest rekening worden gehouden met de beperkte tijd, ongeveer 15 weken.

De afbakening heeft plaatsgevonden op de onderstaande punten:

- **Kwaliteit:** het project moest van voldoende kwaliteit zijn. De kwaliteit was voldoende als:
 - o Documentatie consistentie met elkaar zou zijn.
 - o het systeem uiteindelijk voldoet aan alle wensen van de opdrachtgever zoals deze zijn vastgelegd in het business requirements report;
 - o het systeem betrouwbare uitvoer levert;Het gebouwde systeem moest getest worden om te controleren of de uitvoer juist zou zijn. Door te testen is de kwaliteit te waarborgen. Verder moest er een duidelijke samenhang zijn tussen de ontworpen modellen. Wanneer de opdrachtgever tevreden was kon ik concluderen dat de kwaliteit goed zou zijn.
- **Budget:** niet van toepassing.
- **Mensen en middelen:** omdat het project door een persoon werd uitgevoerd is dit een afbakening. Tevens was het hier van belang dat de belanghebbenden bereid moesten zijn input te leveren aan de uitvoerder van het project.
- **Tijd:** afbakening van tijd is maximaal 18 weken. Uitloop is in principe niet mogelijk. De laatste drie weken zijn gereserveerd voor het opstellen van het procesverslag.
- **Producten:** dit was een van de belangrijkste zaken die afgebakend moest worden. Wat gaan we wel meenemen en wat niet? De opdrachtgever heeft in het voortgangsgesprek in de eerste week aangegeven dat hij ongeveer twee KPI's per proces wilde zien, uitgaande van totaal vier processen. Dit leek mij redelijk. Omdat ik in de eerste week niet wist welke bronsystemen er nodig waren om de gewenste rapportages te verkrijgen heb ik met de opdrachtgever afgesproken om in de vierde week de definitieve afbakening vast te leggen.

5.4 Infrastructuur

Het was van belang even stil te staan bij de capaciteit die nodig zou zijn wanneer de oplossing geïmplementeerd zou zijn. Dit onderwerp is meegenomen bij het gesprek met op de opdrachtgever in de eerste week. De opdrachtgever heeft te weten gegeven dat er geen aanpassingen nodig zijn in de huidige infrastructuur. De server zou krachtig genoeg zijn om de rapportages te generen. Indien het wel nodig zou zijn aanpassingen te doen zou dit geen probleem zijn. In dit gesprek heeft hij ook aangegeven dat er geen historische gegevens opgeslagen hoefden te worden. Het zou voldoende zijn om de gegevens uit het huidige jaar mee te nemen t/m de laatste maand van het jaar om het budget en de forecast tegen elkaar af te zetten.

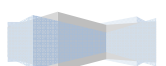


5.5 Personeel en ervaring

Business Intelligence projecten zijn projecten waarbij niet een persoon binnen de organisatie betrokken is maar waar vaak meerdere personen een rol in spelen, zo ook bij dit project. Uiteindelijk zullen er meerdere personen gebruik maken van de oplossing. Het was dus belangrijk om te weten wie er gebruik zouden gaan maken van de oplossing om de juiste informatiebehoefte te kunnen bepalen.

ETTU maakt gebruik van een Microsoft Sharepoint 2007 intranet waarin alle processen een eigen omgeving kennen. In deze omgeving stonden ook de diverse proceseigenaren benoemd. Ik heb in deze omgeving rondgekeken en de proceseigenaren genoteerd. Zodoende wist ik wie de proceseigenaar was en dus ook de persoon die ik moest benaderen in de volgende fase waarin de informatiebehoefte bepaald zou worden. Uiteindelijk waren er vier personen die als belanghebbenden golden voor dit project. Dit waren de proceseigenaren van de processen: bestuur, verkoop, techniek en interne communicatie.

Tevens was het van belang te weten welke personen binnen de organisatie deskundig waren op een bepaald gebied. Dit zou handig kunnen zijn wanneer er hulp nodig was bij het ontwikkelen. Er is in een vroeg stadium contact gelegd met de collega's om te vragen wat voor soort werk zij deden en waar ze goed in waren. De meeste technische kennis die handig kon zijn was te vinden in de businessunit Applications waar de toepassingen werden ontwikkeld. Zij hadden verstand van SQL Server en .NET codes. In de businessunit Configurations was kennis te vinden over Sharepoint 2007. Dit kon handig zijn bij het ontwikkelen van het managementdashboard. Helaas waren er geen Business Intelligence consultant werkzaam bij ETTU. Omdat er voor dit project voornamelijk zou worden gewerkt met Microsoft SQL Server 2005 Business Intelligence Studio moest ik dit mijzelf aanleren. Het was dus nuttig om in vroeg stadium deze applicatie te verkennen en de mogelijkheden te bekijken. In de planning is er geen concreet tijdsbestek gepland om met de tools te experimenteren. Dit is ingecalculeerd in de ontwikkelingfase van het project.



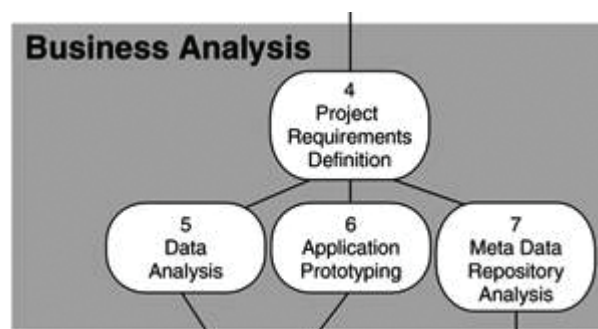
6.0 Fase 3: Business Analysis

6.1 Inleiding

De volgende fase was de business analysis fase. Het belangrijkste binnen deze fase was het bepalen van de informatiebehoefte. Dit was de fundering van het project. Er moest in kaart worden gebracht aan welke informatie men behoefte heeft om vervolgens te kunnen bepalen hoe deze behoefte vervuld kan worden, gebruik makend van de data die op dit moment ligt opgeslagen in de bronsystemen van ETTU.

Deze fase was opgesplitst in de volgende vier deelfasen:

- Project Requirements Definition;
- Data analysis;
- Application prototyping;
- Meta-data repository design.



Figuur 6 Fase 3: business analysis

6.2 Project Requirements opstellen

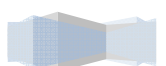
Nu het belang van het project duidelijk was, de proceseigenaren bekend waren en er meer informatie was verkregen over het bedrijf, kon er gestart worden met het opstellen van de project requirements. De output van deze fase was het zogenaamde “business requirements report”. Dit is het onderdeel waarin er is gekeken welke informatiebehoefte er was bij de diverse proceseigenaren en op welke vragen zij geen antwoord konden krijgen.

De hoofdactiviteiten binnen deze fase waren:

- Interviews afleggen voor het bepalen van de informatiebehoefte;
- Prioriteiten toekennen aan de informatiebehoefte (ivm beperkte tijd);
- Data kwaliteit beoordelen (adhv het KING model);
- Opstellen van een business requirements report.

Speerpunten van het Business Requirements Report:

- Informatiebehoefte per proces: welke proceseigenaar wil welke informatie graag zien?
- Succesfactoren en Prestatie-Indicatoren: waarop stuurt ETTU?
- Uitwerking in KPI's: informatie vertalen naar meetbare feiten en dimensies.
- Datakwaliteit: waarop moet men letten als het gaat om datakwaliteit.



Het eerste onderdeel van het rapport geeft een beschrijving van de succesfactoren van ETTU. Normaal gesproken wordt er aan het begin van een Business Intelligence project nagedacht over hoe een organisatie gestuurd wordt, welke succesfactoren er zijn en hoe men de prestaties kan meten met behulp van prestatie-indicatoren. ETTU wist al voor de start van dit project welke KPI's zij in het dashboard wilde hebben en waarop zij wilde sturen. Toch vond ik het nuttig om alles goed om een rijtje te zetten om de situatie beter te begrijpen. Waren het de juiste indicatoren waarop men wilde gaan sturen, waarom wilde men sturen met deze factoren, begreep ik het nut ervan, waren er geen betere stuurfactoren te bedenken? Dit waren de vragen die naar boven schoten, daarom heb ervoor heb gekozen om een hoofdstuk over succesfactoren te schrijven in het business requirements report. Om de juiste informatie over de succesfactoren te vinden heb ik gebruik gemaakt van het businessplan van ETTU. Hierin had men een missiestatement geformuleerd waarin een aantal belangrijke succesfactoren waren geformuleerd. De complete lijst met de 10 algemene succesfactoren is te vinden in bijlage C. Verder had men per bedrijfsproces een aantal succesfactoren geformuleerd. Ik kon deze dus eenvoudig overnemen in mijn rapport. Vervolgens ben ik gaan kijken op welke succesfactoren de gekozen KPI's betrekking hadden. Onderstaande vier succesfactoren wilde men graag meten:

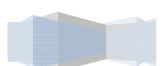
- Projectvolgning;
- Winstgevendheid;
- Klantgerichtheid;
- Inzicht in prospectieve situatie.

Interviewen

Om erachter te komen aan welke informatie men behoefte had heb ik besloten om met iedere proceseigenaar een interview af te leggen. Interviewen was een snelle en effectieve manier om achter de juiste informatie te komen.

In de eerste en tweede week van het project heb ik met iedere proceseigenaar een afspraak gepland in zijn agenda met als onderwerp 'KPI's bepalen'. De proceseigenaren waren bekend met deze term, dat had ik van de opdrachtgever meegekregen. Voordat ik ben begonnen met dit project is er al gesproken over de zogenaamde 'cockpit', het managementdashboard. Ze hadden dus min of meer nagedacht welke informatiebehoefte ze hadden. Hiermee probeer ik te zeggen dat de proceseigenaren niet vreemd waren van waar ik mee bezig was en dat zij zich ook betrokken voelden bij het project, hetgeen in mijn voordeel was.

Als voorbereiding op de interviews heb ik een agenda gemaakt. Deze agenda is ook naar de proceseigenaren gestuurd waarmee het interview afgelegd zou gaan worden. Onderstaand figuur toont een voorbeeldagenda van een interview.





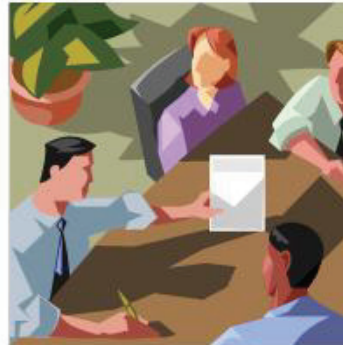
Notulen vergadering

Plaats: Demo

Datum: maandag 12-2-2007

Tijd: 10:50 - 11:30

Aanwezig: Ronald Kraijesteijn, Patrick Smeets



1. Openingen

De vergadering is geopend nadat iedereen aanwezig bleek te zijn.

2.0 Algemeen

Patrick regelt de planning van projecten. Hij is de schakel tussen de klant en de ontwikkelaars. Ettu werkt op basis van een vaste projectprijs die vooraf met de klant wordt afgesproken. Patrick is daarom geïnteresseerd in het volgende:

- Inzicht in de huidige stand mt een project.

Het doel: meer ruimte hebben om in te grijpen. Hoe eerder je iets weet, hoe meer ruimte er is voor maatregelen. Verder kunnen de ontwikkelaars ook de huidige stand van het project zien in geld en uren.

3.0 KPI's

Patrick is geïnteresseerd in onderstaande feiten:

- Factuurabele uren / niet-factuurabele uren;
- Bezettingspercentages;
- Gemiddeld uurtarief;
- Omzet / budget;
- Winst per project;
- Ziekteverzuim.

4.0 Overige

Voorbeeld:

Begroet voor project A: 50.000

Maand 1: 20.000 aan declarabele uren aan dit project gewerkt

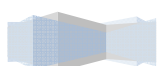
Maand 2: 25.000 aan declarabele uren aan dit project gewerkt

Totaal: 5.000 winst op dit project

Als er nog 5000 aan uren was besteed was dit project break-even. Als er overheen was gegaan was er verlies gedraaid op dit project.

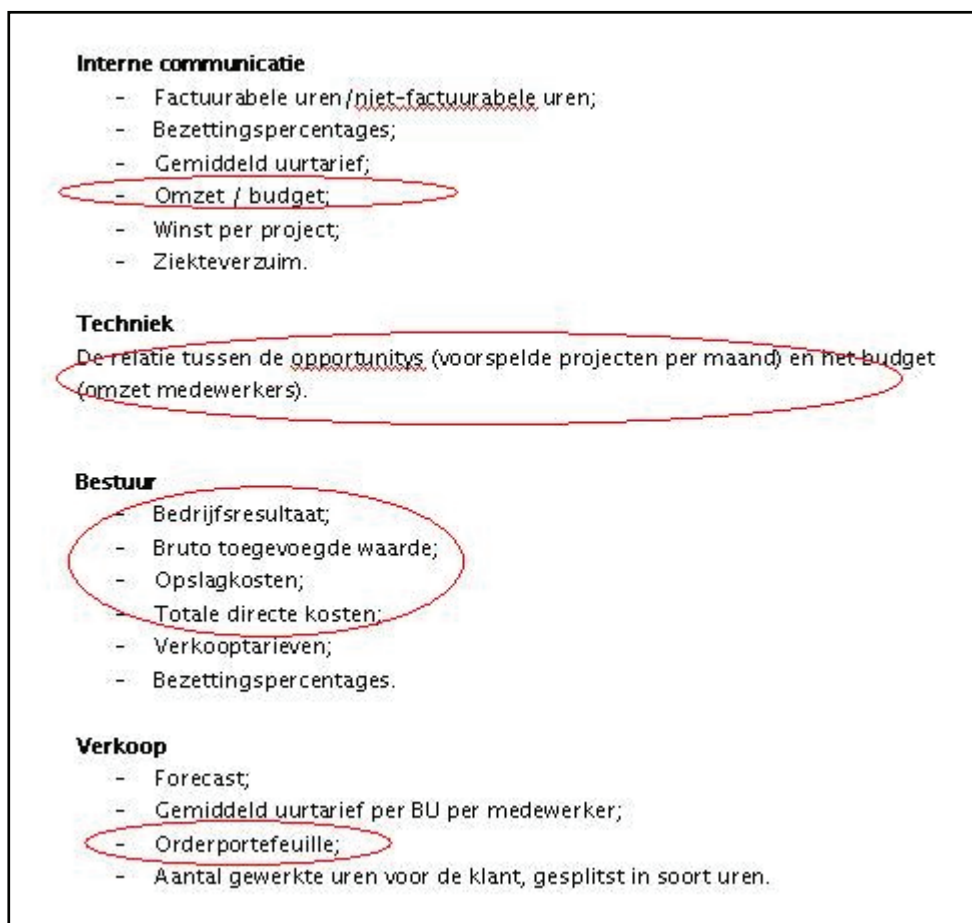
Patrick wil een rond metertje waarop hij kan zien hoe ver een project gevorderd is met de volgende indicatoren: Break even, winst en verlies. Hetzelfde geldt voor de omzet met de volgende indicatoren: 0, break-even, budget, etc.

Figuur 8 Notulen vergadering



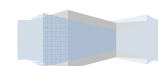
Prioriteiten toekennen aan de informatiebehoefte

Nadat alle interviews waren afgelegd was de informatiebehoefte helder. Omdat deze informatiebehoefte omvangrijk was en niet alles in het project zou passen leek het mij verstandig om een prioriteitenlijst te hanteren. Dit was nodig om te bepalen welke KPI's als eerste uitgewerkt moesten worden en dus het belangrijkste waren voor het management. Om deze lijst te kunnen opstellen moest ik weten welke KPI's men het belangrijkste vond. Dit waren de KPI's die voortvloeide uit de bedrijfsstrategie.



Figuur 9 Prioriteiten informatiebehoefte

Ik heb in week drie een presentatie gehouden voor de proceseigenaren waarin ik de KPI's heb gepresenteerd. Om de presentatie overzichtelijk te houden heb ik de KPI's onderverdeeld naar proces. Zo bestond de presentatie uit vier stukjes, KPI's bestuur, KPI's verkoop, KPI's techniek en KPI's interne communicatie. Aan het einde van de presentatie heb ik gemeld dat het noodzakelijk was om prioriteiten toe te kennen aan de KPI's in verband met de beperkte tijd. Ik heb voorgesteld om vijf KPI's te nemen die we in ieder geval uit zouden gaan werken. Beter vijf goed uitgewerkte KPI's die gepresenteerd worden dan tien die een beetje werken. Mocht er vervolgens tijd over zijn dan kon er altijd nog besloten worden een extra KPI mee te nemen in het project. De proceseigenaren vonden het een goed idee om het zo aan te pakken en hebben hun voorkeur opgegeven. Zo ontstond een lijstje met vijf KPI's. Deze zijn in bovenstaande afbeelding aangegeven met het rode rondje. De proceseigenaar van de interne communicatie heeft te kennen gegeven geïnteresseerd te zijn in budgetverbranding van projecten. Techniek was vooral geïnteresseerd in de verhouding tussen orders en omzet wat ook gold voor de proceseigenaar van de verkoop. Tot slot was de



proceseigenaar van bestuur geïnteresseerd in de financiële gegevens. Omdat ik voor de presentatie al wat kennis had opgedaan over de bronsystemen wist ik dat dit goed te doen zou zijn. Ook zou het project van voldoende omvang zijn omdat ik de data uit drie verschillende bronsystemen moest halen, hetgeen vrij lastig te realiseren zou zijn. Ik ben akkoord gegaan met de keuze van de proceseigenaren.

Onderstaand figuur toont alle potentiële KPI's met het bijhorende proces. Dit zijn alle KPI's waarin men geïnteresseerd was alvorens er prioriteiten waren gegeven aan de KPI's. Verder is er in dit overzicht af te lezen bij welk proces de KPI hoorde en welke proceseigenaar geïnteresseerd was in welke KPI. Dit zou de opdrachtgever een beeld geven van welke zaken gemeten zouden gaan worden en welke rechten er uitgedeeld moesten worden aan het dashboard. Het zou mogelijk kunnen zijn dat niet iedere gebruiker van het dashboard toegang zou hebben tot alle gegevens, vandaar dat deze kolom is opgenomen.

KPI	Proces	Weergave
Opslagkosten (maand,jaar,BU,totaal)*	Bestuur	Piet
Bruto toegevoegde waarde (maand, BU, totaal)*	Bestuur	Piet
Bedrijfsresultaat (maand, BU, totaal)*	Bestuur	Piet
Directe kosten per maand (BU, totaal)	Bestuur	Piet
Winst per project	Techniek	Patrick, Eric-Jan, Piet
Percentage gebruikt budget per project	Techniek	Patrick / Piet
Ziektepercentage (medewerker,BU,jaar,maand)	Techniek	Piet, Patrick
Verlofpercentage (medewerker, BU)	Techniek	Piet
Gemiddelde omzet per klant	Verkoop	Piet / Vincent
Gemiddeld verkooptarief (maand, jaar, BU, medewerker)*	Verkoop	Piet, Patrick, Vincent
Bezettingsgraad (BU, medewerker)	Verkoop	Piet, Patrick
Orderportefeuille per maand*	Verkoop	Piet, Eric-Jan,Vincent
Werkelijke omzet*	Verkoop	Patrick
Aantal nieuwe klanten per maand	Verkoop	Piet, Vincent

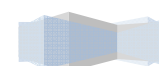
***Notities**
Opslagkosten per maand: als percentage vd bruto toegevoegde waarde
Bruto toegevoegde waarde: afgezet tegen het budget
Bedrijfsresultaat: afgezet tegen het budget
Directe kosten per maand: als percentage vd bruto toegevoegde waarde
Gemiddeld verkooptarief: afgezet tegen het budget
Orderportefeuille: als percentage vd omzet uit de forecast
Werkelijke omzet: als percentage van het budget

Figuur 10 Overzicht KPI's

Uitwerking van de KPI's

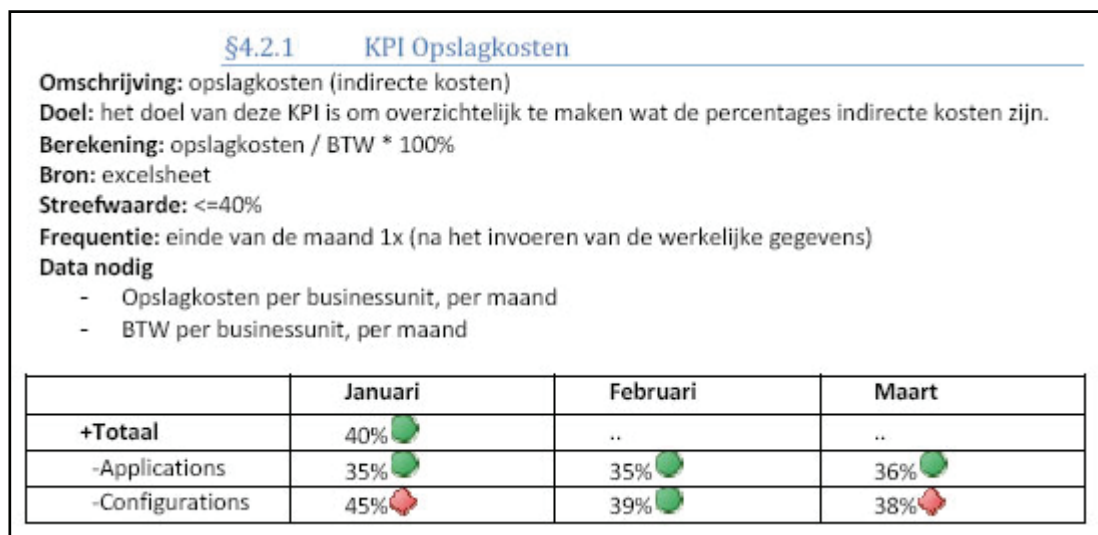
De volgende stap was het uitwerken van alle KPI's. Er is gekozen om de KPI's uit te werken naar proces, mede omdat ETTU aangegeven heeft dit graag te willen. Dit hoofdstuk geeft voor alle KPI's een aantal zaken weer. De lijst met de punten die uitgewerkt moesten worden is niet afkomstig uit een boek of een methode, maar is door mijzelf opgesteld. De opdrachtgever heeft aangegeven een overzicht te willen hebben van de uitgewerkte KPI's. De methode die is gebruikt geeft niet aan hoe dit gedaan moest worden. Ik moest dus nadenken over welke punten ik hierin op zou gaan nemen. Ik heb nagedacht over welke gegevens relevant konden zijn voor de opdrachtgever en welke gegevens voor mijzelf interessant konden zijn bij de uitwerking van de KPI's. Op deze manier is de lijst ontstaan:

- Omschrijving;
- Doel;
- Berekening;



- Bron;
- Streefwaarde;
- Frequentie;
- Benodigde data;
- Uitvoer.

Bij de omschrijving is een korte omschrijving te vinden van de KPI. Het doel beschrijft waarom het nuttig is om deze KPI weer te geven en welke informatie inzichtelijk wordt gemaakt. De berekening laat zien hoe deze KPI berekend kan worden en de bron toont aan uit welke bronnen de gegevens afkomstig zijn die nodig zijn om de KPI weer te geven. De streefwaarde geeft een waarde op waarbij de KPI een “groen licht” geeft. Er wordt dus gestreefd om onder deze waarde te blijven. De frequentie laat zien hoe vaak de KPI verversd dient te worden. Dit is afhankelijk van de brondata. Uurgegevens worden bijvoorbeeld slechts een maal per week ingevoerd, deze KPI hoeft dan ook maar eens per week verversd te worden. Het veld ‘benodigde data’ geeft aan welke data nodig is uit de bronsystemen en het veld ‘uitvoer’ bevat een grafische weergave van hoe het dashboard eruit moet komen te zien voor deze KPI. Zoals in onderstaande afbeelding te zien is zullen de indirecte kosten per maand worden getoond en ook per businessunit.



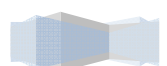
Figuur 11 Uitgewerkte KPI

Bovenstaand plaatje is voor iedere KPI uitgewerkt. De volgende stap in het project was het inspecteren van de bronsystemen en de brondata, ook wel data-analyse genaamd.

6.3 Data-analyse

De volgende stap in mijn project was het uitvoeren van een data-analyse. Om een Business Intelligence oplossing te maken wordt er gebruik gemaakt van data uit diverse bronsystemen. De data-analyse was nodig om te weten te komen welke data er nodig zou zijn om een antwoord te kunnen geven op de managementvragen die eerder zijn besproken. Door het uitvoeren van de data-analyse heb ik geprobeerd antwoord te krijgen op de volgende vragen:

- Welke databronnen worden er gebruikt;
- Welke databases worden er gebruikt;
- Welke tabellen worden er gebruikt;
- Welke gegevens worden er gebruikt;



- Wat is dat kwaliteit van de data.

Ik vond het belangrijk om in een vroeg stadium van het project te kijken naar de brondata. Dit zou handig zijn om het project goed in te kunnen schatten en op tijd melding te kunnen maken van problemen. Mocht blijken dat de data slecht was opgeslagen of niet werd bijgehouden kon dit gemeld worden om zo toch de gewenste rapportages te verkrijgen. Mocht blijken dat er veel systemen werden gebruikt en grote databases in gebruik waren, kon er worden besloten het project meer af te bakenen. Mijn aanpak in dit stadium was om per systeem te kijken hoe het systeem werkte en hoe de data was opgeslagen. Om dit per systeem te bekijken moest ik als eerste bepalen welke systemen relevant waren voor mijn project.

Data-analyse budgetteringssysteem

Ik was al op de hoogte van het feit dat de budgetgegevens in een budgetsysteem werden bijgehouden, dat had ik tijdens de bespreking met de proceseigenaar van het proces 'bestuur' al te horen gekregen. Ik ben dan ook begonnen om naar dit systeem te kijken omdat ik al de nodige informatie over de werking van dit systeem had gekregen.

Om beter te begrijpen hoe ik naar het systeem heb gekeken is het noodzakelijk om wat meer informatie te geven over de opbouw van het systeem. Het budgetsysteem was een excelsheet dat bestond uit ongeveer vijftien tabbladen. Er waren een aantal tabbladen waarin het budget werd beschreven en een aantal tabbladen waarin de forecast werd beschreven. Forecast lijkt erg op budget, alleen worden hier ook de werkelijke cijfers in beschreven. Per businessunit was er dus logischerwijs een tabblad "budget" en een tabblad "forecast" aanwezig. Onderstaande tabel toont een overzicht van hoe het systeem was opgebouwd. Zo hebben we per businessunit een forecast en een budget.

Forecast applications	Budget applications	Forecast configurations	Budget configurations	Consolidatie forecast	Consolidatie budget
-----------------------	---------------------	-------------------------	-----------------------	-----------------------	---------------------

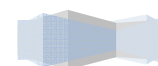
De laatste twee kolommen zijn de samengevatte gegevens van de tabbladen forecast en budget. In eerste instantie was de excelsheet erg onduidelijk. Er stonden veel gegevens op waardoor niet bekend was waar de data stond die nodig was. De proceseigenaar van het proces bestuur had uitgelegd hoe het ongeveer werkte maar nu ik het systeem voor mij had begreep ik niet hoe het precies in elkaar zat. Door het bestuderen van de tabbladen begreep ik na een tijdje hoe het systeem werkte. Onduidelijkheden zijn genoteerd en teruggekoppeld aan de proceseigenaar. Nu moest ik gaan bepalen waar exact de gegevens stonden die ik nodig zou hebben voor het kunnen beantwoorden van de managementvragen.

De proceseigenaar van het proces 'bestuur' wilde graag de volgende KPI's in zijn managementdashboard:

- Bruto toegevoegde waarde;
- Directe kosten;
- Indirecte kosten.

Het was duidelijk dat al deze gegevens in het budgetteringssysteem waren vastgelegd, alleen waren deze gegevens te vinden in meerdere tabbladen. Het was op dat moment nog niet duidelijk hoe hij de gegevens wilde hebben, bijvoorbeeld in geldbedragen of percentages. Om dit helder te krijgen heb ik een korte meeting gepland met de proceseigenaar. Daaruit heb ik de volgende conclusies kunnen trekken:

- Bruto toegevoegde waarde in geld, afgezet tegen het budget;



- Directe kosten als percentage van de bruto toegevoegde waarde, afgezet tegen het budget;
- Indirecte kosten als percentage van de bruto toegevoegde waarde, afgezet tegen het budget;

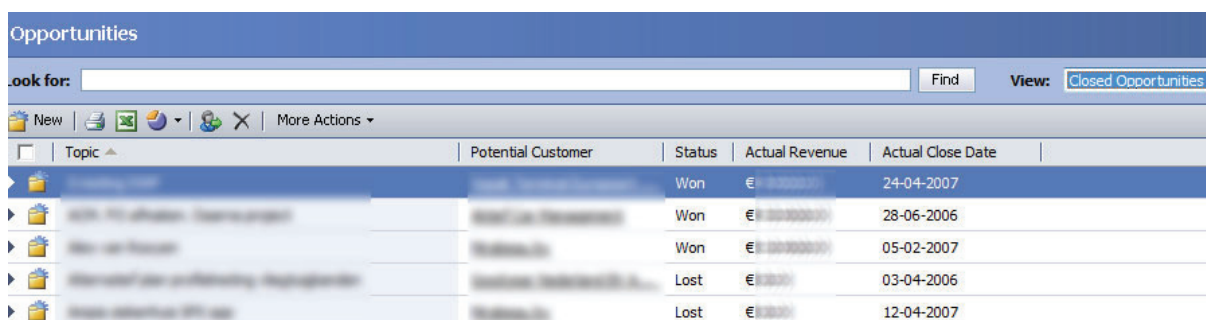
Nu ik dit wist kon ik gaan kijken op welke tabbladen de gegevens stonden. Het ging bij alle drie de vragen om het vergelijken van het budget met de forecast. Ik moest dus als eerste een getal halen uit de forecast en daarna uit het budget om zo de gegevens tegen elkaar af te kunnen zetten. Op dit moment wist ik voldoende. Ik had een beter beeld van hoe het systeem werkte en waar ik de gegevens kon vinden die ik nodig had. Bij de volgende fase, het design, zal ik meer vertellen over welke waarden ik heb gebruikt.

Data-analyse CRM-systeem

Bij het bepalen van de informatiebehoefte was naar voren gekomen dat de proceseigenaar van het proces verkopen behoefte had om informatie te verkrijgen over de orderportefeuille. Hij heeft aangegeven orders op te slaan in een CRM-systeem. Het CRM systeem was het systeem dat verantwoordelijk was voor de beheersing van alle relaties van ETTU. Hierin stonden bijvoorbeeld de orders van klanten en dus ook de orderbedragen.

Om meer gevoel te krijgen van de werking van het systeem leek het mij een goed idee om eens te kijken hoe men het systeem gebruikte. Ik heb aan de proceseigenaar gevraagd of hij wilde demonstreren hoe hij orders invoerde. Hij heeft vervolgens laten zien hoe hij de orders invoerde en bijwerkte. Ter verduidelijking zal ik een omschrijving geven van de werking van het systeem zoals de proceseigenaar mij dit ook heeft uitgelegd.

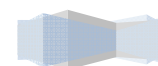
Een order is een project of een aantal projecturen dat is verkocht maar het kan ook een potentieel project zijn dat nog moeten worden verkocht. Zodra een verkoopmedewerker van ETTU een eerste klantencontact heeft gelegd met een potentiële klant is er een kans dat dit een order gaat worden. Er wordt ook direct een schatting gemaakt van hoeveel het project gaat opleveren en wat de kans is dat de order doorgaat. Bij een eerste klantencontact is deze kans meestal 10%. Hoe meer contact er met een klant is, hoe duidelijker wordt dat een order wel of niet doorgaat. Dit percentage wordt dan ook regelmatig aangepast. Gaat een order niet door dan wordt de status "lost" meegegeven. Gaat de order wel door en de kans dus 100% is zal de status veranderen in "won". Bij het invoeren van de order wordt een belangrijke parameter meegegeven, namelijk de datum dat de order gaat vallen. Onderstaand figuur toont een screenshot van het systeem om een idee te geven hoe dit eruit ziet.



Topic	Potential Customer	Status	Actual Revenue	Actual Close Date
Topic 1	Potential Customer 1	Won	€ 100.000,00	24-04-2007
Topic 2	Potential Customer 2	Won	€ 100.000,00	28-06-2006
Topic 3	Potential Customer 3	Won	€ 100.000,00	05-02-2007
Topic 4	Potential Customer 4	Lost	€ 100,00	03-04-2006
Topic 5	Potential Customer 5	Lost	€ 100,00	12-04-2007

Figuur 12 Opportunities in CRM

Na de demonstratie van de werking van het systeem heb ik een kopie van de originele database laten maken door een technisch medewerker en heb ik deze geïmporteerd in mijn SQL server in mijn virtuele werkomgeving. Zo kon ik de database bekijken zonder dat de live-omgeving gevaar liep.



Nadat de database geïmporteerd was ben ik gaan zoeken naar de tabellen waarin de ordergegevens stonden. Ik moest namelijk weten in welke tabel de gegevens waren opgeslagen die in het ordersysteem werden gebruikt bij het onderdeel opportuniteiten. Omdat de database uit meer dan 100 tabellen bestond leek het mij verstandig om internet te raadplegen. De eerste site waar ik op heb gekeken was die van de fabrikant, Microsoft. Op de website van Microsoft was een zogeheten SDK (software development kit) beschikbaar waarin veel informatie te vinden was over het systeem. Ik kwam er op deze manier achter in welke tabel de gegevens waren opgeslagen, namelijk de tabel: OpportunityBaseCode. Om erachter te komen of dit de enige tabel was die ik nodig had heb ik een test uitgevoerd. Ik wilde graag weten of ik de som van alle orders in een bepaalde maand tevoorschijn kon krijgen. Ik heb dit getest door een SQL-query op de database los te laten binnen de Microsoft SQL Server Management Studio, een tool die gebruikt wordt om de databases te bekijken. Na een aantal pogingen lukte het om deze query uit te voeren. Ik heb ook geprobeerd om de data tevoorschijn te halen uit bovenstaande screenshot. Dit is gelukt door gebruik te maken van deze ene tabel. Voor mij was het nu duidelijk dat ik slechts een tabel nodig had uit deze grote database.

Data-analyse urensysteem

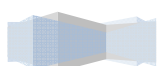
De proceseigenaar van de interne communicatie heeft te kennen gegeven geïnteresseerd te zijn in de budgetverbranding van projecten. Om het gebruikte budget per project te kunnen uitrekenen had ik informatie nodig over projecten, uren en uurtarieven. Tijdens een van de gesprekken met de opdrachtgever heb ik gevraagd hoe deze uren bijgehouden worden. Er bleek een applicatie op het netwerk te draaien in de vorm van een webapplicatie waarin de uren bijgehouden worden. Iedere medewerker moest aan het einde van de maand zijn gewerkte uren invoeren in het systeem. Omdat ik graag wilde weten hoe het systeem werkte en hoe de database opgebouwd was heb ik een logincode van het systeem gevraagd en een kopie van de originele database. Deze database bleek een MySQL database te zijn.

Voordat ik ben begonnen met het analyseren van de database leek het mij verstandig eerst te kijken hoe het programma werkte waarin de uren werden vastgelegd. Ik heb aan een willekeurige collega (die het systeem gebruikt) gevraagd of hij kon demonstreren hoe hij zijn uren invulde. Dit bleek relatief eenvoudig te zijn. Het enige dat hij hoefde in te vullen waren de juiste uren bij het juiste project. Niet declarabele uren werden op een intern project geschreven of een project “ziekte of verlof”.

Hierna ben ik met mijn eigen inlogcode met beheerderbevoegdheden ingelogd in het systeem. Voor het project moest ik op zoek gaan naar de volgende meetwaarden:

- Afgebakend budget/aantal uren voor een project;
- Koppeling tussen medewerkers en projecten;
- Uurlonen per project.

Deze meetwaarden waren nodig om de juiste informatie te kunnen weergeven in het managementdashboard.

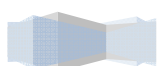




Figuur 13 Projecten in het urensysteem

Ik heb de applicatie bekeken die niet erg moeilijk te doorgronden was. Ik had snel gevonden waarnaar ik op zoek was, namelijk de bovenstaande drie onderdelen. Er was een onderdeel binnen de applicatie genaamd “beheer” waarin de klanten gedefinieerd waren. Aan iedere klant waren een of meerde projecten gekoppeld. Ook was per project vastgelegd welke medewerkers uren mochten schrijven voor dit project, wat de afbakening van het project was in uren/budget en wat het uurloon binnen een project was, precies wat ik zocht. Er was tevens een koppeling tussen project en medewerker aanwezig waar weer een ander uurloon voor stond. Het was even onduidelijk of ik naar het uurloon voor het project moest kijken of naar het uurloon van de medewerker die aan het project was gekoppeld. Ik heb dit even nagevraagd en ik moest naar de tweede optie kijken.

Nu ik wist hoe de applicatie ongeveer werkte was de volgende stap om de database te bekijken. Deze Mysql database was inmiddels geïnstalleerd op mijn virtuele server en was benaderbaar vanaf mijn browser door middel van PHPmyadmin. Ik ben ingelogd en heb een blik geworpen op de tabellen. De database bestond uit ongeveer tien verschillende tabellen. De tabellen waarnaar ik op zoek was moesten gegevens bevatten over de projecten en de geschreven uren per project. Een tabel had de logische naam “uren”. Hierin vond ik alle geschreven uren, alleen was voor mij de koppeling met een andere tabel die ik verwachtte, namelijk de projectentabel, onvindbaar. Het probleem was dat de sleutels die in de verschillende tabellen werden gebruikt verschillende hadden waardoor de relaties moeilijk te leggen waren. Ook is er binnen Mysql geen mogelijkheid om een relatieschema van de tabellen op te vragen. Het was voor mij een raadsel hoe de tabellen met elkaar verbonden waren. De enige manier om hierachter te komen was om de fabrikant van het urensysteem, genaamd “Clockwise” een e-mail te sturen.



De e-mail naar de fabrikant heb ik ingeleid door een korte situatieschets. Dit leek mij handig omdat men anders misschien niet mee wilde werken. Vervolgens heb ik het probleem uitgelegd. Het probleem was eigenlijk dat ik de relaties niet goed kon zien tussen de twee tabellen die ik nodig had, “uren” en “projecten”. Ik heb daar een voorbeeld bij gegeven zodat het helder was. Ongeveer twee dagen later had ik al een antwoord met een uitleg van hoe de relaties in elkaar zaten. Helaas was het voor mij nog niet helemaal duidelijk. Ik heb daarom hulp ingeschakeld van een collega die expert was op het gebied van databases. Hij heeft naar de e-mail gekeken en naar de tabellen. Binnen enkele minuten had hij door hoe de relaties in elkaar zaten. De moeilijkheid zat hem in het volgende; een project kon bestaan uit meerdere subprojecten die vervolgens in dezelfde tabel waren ondergebracht waardoor het onoverzichtelijk was welke sleutels er gebruikt werden. Ook stonden de gekoppelde medewerkers in dezelfde tabel als de projecten. Ieder project had een ID. Iedere medewerker ook, alleen was hiervoor een extra key opgenomen, de PAR. Had een medewerker ID 100 en PAR99 dan was hij gekoppeld aan een project met ID 99. Iemand als ik met weinig ervaring op databasegebied vond dit eerst lastig te begrijpen. Naderhand was het vrij logisch, alleen moest je het wel even zien.

U_structuur			
ID	PAR	NAAM	MEDEWERKERID
1	1	Project A	NULL
2	1	Subproject B	NULL
3	2	NULL	12
4	2	NULL	13

Zoals in bovenstaande tabel is af te lezen is Project A aanwezig met subproject B waaraan twee medewerkers zijn gekoppeld, 12 en 13.

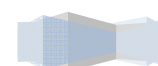
Datakwaliteit

Om te komen tot de juiste rapportages en KPI's is het van essentieel belang dat de brondata van een juiste kwaliteit is. Ik heb hoofdstuk vijf van het rapport gebruikt om dit belangrijke aspect van Business Intelligence te behandelen. Onjuiste data betekent slechte beslissingen. Het doel van dit hoofdstuk was ten eerste de opdrachtgever bewust maken van de essentie van juiste data. Hij is uiteindelijk verantwoordelijk voor de juiste werking van het systeem. Daarnaast is er gekeken welke gevaren er scholen binnen de huidige systemen, lettende op datakwaliteit. Ik heb dit getracht weer te geven zodat de opdrachtgever wist waar hij op moest letten voor de juiste werking van het systeem.

Als eerste is er een bronnenmodel opgesteld zodat voldoende duidelijk is welke bronnen gebruikt worden voor het project. Het beoordelen van de datakwaliteit is dan ook van toepassing op alle bronnen uit het bronnenmodel.

Om de datakwaliteit te beoordelen moest ik eerst weten op welke punten datakwaliteit te beoordelen is. Al snel bleek dat er gigantisch veel informatie te vinden was over kwaliteit en testen van data. Op de website van testnet.org vond ik een artikel over het KING model om datakwaliteit te beoordelen. Dit model bestond uit vijf hoofdcategorieën met daaronder een aantal punten waarop beoordeeld kon worden. Ik heb ervoor gekozen om enkel de punten uit het KING model te behandelen die van toepassing waren op de bronsystemen van ETTU. Het KING model spreekt bijvoorbeeld over punten als versleuteling, traceerbaarheid, zeldzaamheid, etc. Dit soort zaken waren niet van toepassing op het project. Ik heb daarom besloten om de data te beoordelen op de belangrijkste punten die hieronder staan beschreven:

- Integriteit: juistheid van de data.



- Tijdigheid: beschikbaarheid op het juiste moment.
- Volledigheid: compleetheid van de data.
- Consistentheid: komen gegevens overeen met elkaar.

Later kwam ik erachter dat ik beter de term “redundantie” had kunnen gebruiken in plaats van consistentie omdat ik dat er eigenlijk mee bedoelde.

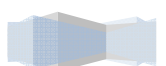
Onderstaand voorbeeld toont de uitwerking van de datakwaliteit van het budgetsysteem.

Integriteit
De proceseigenaar van het proces 'bestuur' zal er op moeten letten dat de data die hij invult integer is. Hij zal dus bij het invullen van het budget/forecast alert moeten zijn op het feit dat er fouten kunnen ontstaan in de rapportages wanneer de data die hij invult niet klopt.
Tijdigheid
Er kan worden uitgegaan dat de excelsheet altijd beschikbaar is. Deze zal worden opgeslagen op het netwerk op dezelfde server als waar het datawarehouse draait. Men zal er voor moeten zorgen dat deze file op een en dezelfde locatie wordt opgeslagen.
Volledigheid
Het is noodzakelijke dat alle velden die nodig zijn om de gewenste rapportage uit te draaien ingevuld zijn. De volgende velden dienen ten minste ingevuld te zijn:
<ul style="list-style-type: none">- Forecast BTW alle maanden, beide businessunits- Forecast opslagkosten, alle maanden, beide businessunits- Forecast directe kosten, alle maanden, beide businessunits
Consistentheid
Niet van toepassing

Figuur 14 Datakwaliteit

Per bronsysteem is er gekeken naar de kwaliteit van de data voor de vier hierboven genoemde punten. Bovenstaand figuur toont een voorbeeld van de kwaliteit van het budgetsysteem. Tevens zijn hieronder nog enkele voorbeelden te vinden:

- De uren moeten voor het weekend in het urensysteem zijn ingevuld omdat deze data ieder weekend wordt ververs in het datawarehouse (tijdigheid).
- De KPI die het gebruikte budget voor een project afzet tegen de projectprijs is alleen zinvol als de projectprijs is ingevuld (volledigheid).
- Sommige projecten zijn dubbel aangemaakt, hier moet op gelet worden (consistentheid).



Opstellen business requirements report

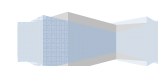
Alle eerder uitgevoerde activiteiten uit dit hoofdstuk zijn verwerkt in het business requirements report die bestond uit onderstaande hoofdstukken.

1.0	INLEIDING	3
2.0	INFORMATIEBEHOEFTE PER PROCES	4
§2.1	Inleiding.....	4
§2.2	Informatiebehoefte proces verkoop.....	4
§2.3	Informatiebehoefte proces bestuur	4
§2.4	Informatiebehoefte proces techniek.....	4
3.0	PRESTATIE-INDICATOREN EN SUCCESFACTOREN ETTU	5
§3.1	Inleiding.....	5
§3.2	Strategische doelen	5
§3.3	Kritische succesfactoren ETTU	5
§3.4	Keuze voor uitwerking Key Succes Factoren	6
§3.5	KPI's per proces.....	7
4.0	UITWERKEN VAN DE KPI'S	9
§4.1	Inleiding.....	9
§4.2	KPI's proces 'bestuur'	9
§4.2.1	KPI Opslagkosten.....	9
§4.2.1	KPI directe kosten	10
§4.2.1	KPI bruto toegevoegde waarde	10
§4.3	KPI's proces 'techniek'	11
§4.3.1	KPI verbruikt budget per project	11
§4.4	KPI's proces 'verkoop'	11
§4.4.1	KPI prospectieve situatie.....	11
5.0	DATAKWALITEIT.....	13
§5.1	Inleiding.....	13
§5.2	Bronsystemen	13
§5.3	Datakwaliteit Excel.....	14
§5.4	Datakwaliteit Clockwise	14
§5.5	Datakwaliteit CRM	15

Figuur 15 Business Requirements Report

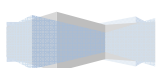
6.4 Application prototyping

Een prototype is een handig hulpmiddel om te kijken of het project voldoet aan alle eisen die de business verwacht. Het is verstandig om een prototype aan het begin van een (groot) project te ontwikkelen zodat missende functionaliteiten al vroeg worden opgemerkt en het project kan worden aangepast. Omdat het hier om een relatief klein project ging met een korte doorlooptijd en weinig ervaring van de ontwikkelaar is er besloten geen prototype te bouwen. Er was onvoldoende kennis van de ontwikkeltools om een prototype al zo vroeg in het project te ontwikkelen. Alvorens deze gebouwd kon worden zou de tijd al om zijn. De methode legt ook uit dat het bouwen van een prototype geen noodzaak is maar alleen moet worden gemaakt indien de tijd en het budget dit toelaat, vandaar dat er is besloten dit onderdeel achterwege te laten. Er is besloten om na een aantal weken een vergadering te houden waarin de (voorlopige) versie van het project zal worden getoond. De opdrachtgever heeft dan nog ruimte om verandering door te geven zodat het project nog op tijd kan worden aangepast.



6.5 Meta-data repository design

Meta-data wordt gebruikt om betekenis te geven aan data. Omdat er zo weinig data wordt gebruikt bij dit project wordt er geen gebruik gemaakt van een meta-data repository. Er zal wel een duidelijke uitleg worden geleverd met de betekenis van de database en alle andere onderdelen van het project.



7.0 Fase 4: Design datawarehouse, ETL-schema's, dashboard

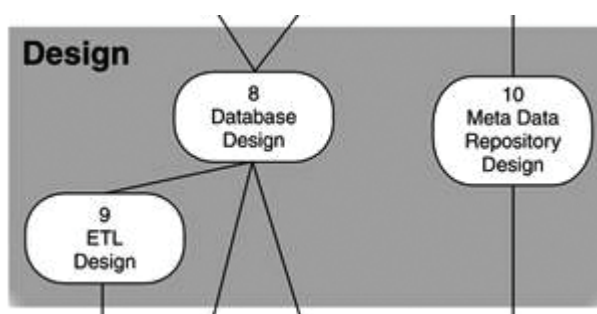
7.1 Inleiding

De volgende fase tijdens de uitvoering van het project was de designfase. De designfase bestond uit het ontwerpen van verschillende modellen die noodzakelijk waren voor een goed werkende Business Intelligence oplossing. De output van deze fase was het design report waarin alle ontwerpen terug te vinden zijn. Deze is te vinden in bijlage D.

Zoals in onderstaand schema is te zien schrijft de gebruikte methode drie deelstappen voor. Dit hoofdstuk beschrijft de belangrijkste activiteiten die in deze fase zijn uitgevoerd. Bij het maken van de ontwerpen zijn de volgende stappen uitgevoerd:

- Opstellen bronnenmodel;
- Opstellen gegevensmodel;
- Opstellen feitenmodel;
- Opstellen sterschema;
- Opstellen ETL-schema's.

Dit is een eigen methode die niet in het boek terug te vinden is. Door deze stappen te volgen ontstond het business design report, uitput van de designfase.

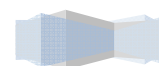


Figuur 16 Fase 4: Design

7.2 Opstellen bronnenmodel

Om een overall beeld te krijgen van alle databases en andere bronnen die gebruikt zouden worden voor het project is er een bronnenmodel opgesteld. Omdat al bekend was welke drie bronnen gebruikt zouden worden kon dit model eenvoudig worden opgesteld.

Het model bevat een overzicht van alle benodigde bronnen maar bevat ook een overzicht van de gegevensgebieden die overlappend zijn tussen de systemen. Zo is uit dit model af te lezen dat zowel het urensysteem als het budgetsysteem gebruik maken van personeelsgegevens. Dit is belangrijk om te weten omdat het te bouwen systeem geen redundante gegevens mag bevatten. Er zal dus in een later stadium worden gekeken welke overlappende gegevens er zijn tussen de systemen. Wanneer deze overlappende gegevens ook worden overgezet naar het datawarehouse zullen hier de nodige bewerking op uitgevoerd (kunnen) worden. Hier wordt later op teruggekomen. Na het opstellen van het bronnenmodel was het noodzakelijk om de opslagstructuur van de benodigde data in kaart te brengen. Dit is gedaan door het opstellen van gegevensmodellen van alle benodigde brondata.



7.3 Opstellen gegevensmodel

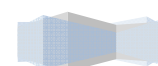
Een gegevensmodel toont zoals de naam al doet vermoeden de gegevens van een bepaalde database of andere bron in een model dat voor iedereen te begrijpen is. Het is noodzakelijk om een gegevensmodel te maken omdat met behulp van dit model duidelijk wordt welke gegevens in welke tabel liggen opgeslagen en in welk formaat deze gegevens liggen opgeslagen. Wanneer er later een ETL-schema opgesteld wordt zal er worden gekeken naar het gegevensmodel om te kunnen bepalen welke data (eventueel) getransformeerd moet worden.

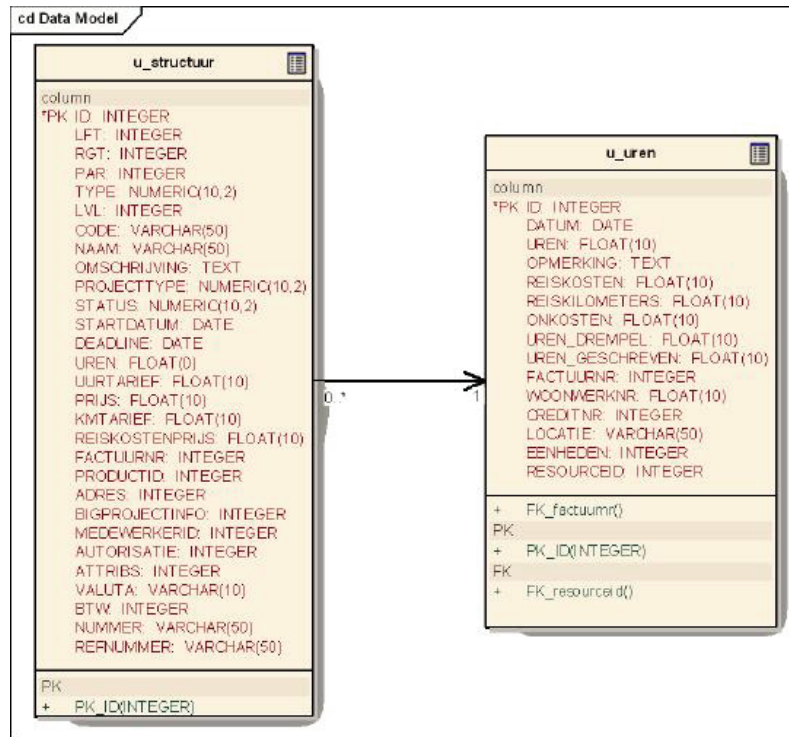
Het opstellen van de gegevensmodellen is gedaan door de opbouw van de tabel of Excelsheet in kaart te brengen. Dit was dus het overnemen van de logische opbouw van de bronsystemen. Het eerste gegevensmodel was die van het budgetstelsel welke in Excel was vastgelegd. Dit stelsel bestond uit een aantal tabbladen waarin de nodige gegevens waren vastgelegd. Per businessunit was er een tabblad aanwezig voor zowel de budget- als de forecastgegevens. Een gegevensmodel bevat enkel de gegevens die nodig zijn voor de ontwikkeling van het datawarehouse. Het was dus niet zinvol om alle gegevens te beschrijven die in de Excelsheet waren vastgelegd maar enkel de gegevens die benodigd waren. Het gegevensmodel voor het budgetstelsel is opgedeeld in zes tabellen die overeenkomen met de zes tabbladen die gebruikt worden. Hieronder zijn de eerste drie te vinden:

Bestand Forecast applications	Attribuut Maand Subtotaal opslagkosten BTW Totale Dir Pers kstn	DataType String Integer Integer Integer	Omschrijving Maand Som kosten Bruto Toegevoegde waarde Directe kosten
Bestand Forecast configurations	Attribuut Maand Subtotaal opslagkosten BTW Totale Dir Pers kstn	DataType String Integer Integer Integer	Omschrijving Maand Som kosten Bruto Toegevoegde waarde Directe kosten
Bestand Forecast detachering	Attribuut Maand Subtotaal opslagkosten BTW Totale Dir Pers kstn	DataType String Integer Integer Integer	Omschrijving Maand Som kosten Bruto Toegevoegde waarde Directe kosten

Figuur 17 Gegevensmodel budgetstelsel

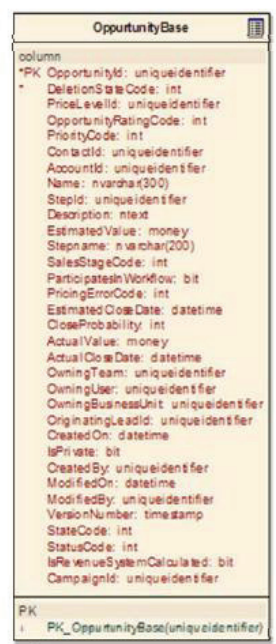
Het volgende gegevensmodel was die van het urenstelsel, de MYSQL database. Ook hier zijn alleen de tabellen gemodelleerd die daadwerkelijk benodigd waren. Het was niet zinvol om de gehele database te beschrijven omdat er slechts een beperkt aantal tabellen nodig waren voor het datawarehouse. Om het gegevensmodel te kunnen maken heb ik een dumpfile gemaakt van de database. Dit is een bestand waarin de structuur van de database terug te vinden is en de gehele inhoud van de tabellen. Zo kon ik op een eenvoudige wijze zien hoe de tabellen waren opgemaakt en welke relaties er waren tussen de tabellen. Onderstaande figuur toont het gegevensmodel van het urenstelsel.



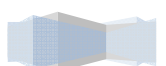


Figuur 18 Gegevensmodel urensysteem

Het laatste model was het gegevensmodel voor het CRM systeem, de grootste SQL-database. Ook van dit systeem was inmiddels bekend welke data benodigd was voor het project. Deze data die nodig was voor het project was vastgelegd in slechts een tabel, daarom bestaat het gegevensmodel uit slechts een tabel. Deze kon opgesteld worden door een connectie aan te maken met de SQL-database en de kolommen over te nemen. Onderstaande afbeelding toont het gegevensmodel van het CRM-systeem.



Figuur 19 Gegevensmodel CRM



Nu alle data die benodigd was voor het datawarehouse in kaart was gebracht kon er gestart worden met het beschrijven van de diverse feiten en dimensies.

7.4 Opstellen feitmodel

Het volgende onderdeel van het designrapport was het feitmodel. De feiten die hierin beschreven staan zijn de feiten die gemeten gaan worden en die in het dashboard getoond moesten worden. Verder toont het feitenmodel de relaties tussen de feiten die gemeten worden en de dimensies in welke deze worden weergegeven. Tot slot is er een kolom in dit model opgenomen waarin de brontabel beschreven staat zodat in een oogopslag te zien is waar de huidige brondata opgeslagen ligt.

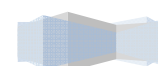
Het model is ontstaan door na te denken over welke zaken gemeten moesten worden. In een vroeg stadium van dit project is er overleg geweest met de diverse proceseigenaren. Het doel hiervan was om de informatiebehoefte helder te krijgen met betrekking tot het project, het maken van een managementdashboard. Daarin kwamen diverse zaken naar voren die men graag wilde gaan meten zoals de projectvoortgang en de orderportefeuille. Het was nu zaak om deze informatiebehoefte te gaan vertalen in meetbare zaken, ook wel feiten genoemd.

Door te kijken naar de informatiebehoefte en het logisch redeneren is de volgende feitentabel ontstaan met meetwaarden:

Feiten
Forecast directe kosten
Forecast indirecte kosten
Forecast Bruto toegevoegde waarde
Budget directe kosten
Budget indirecte kosten
Budget Bruto toegevoegde waarde
Forecast Ebit
Budget Ebit
Verbrand budget
Projectprijs
Omzet Directe Medewerkers
Orderportefeuille

n.b. EBIT betekent voor ETTU "winst".

Deze feiten kwamen overeen met de informatiebehoefte en de onderzoeksvragen die eerder zijn opgesteld. Nu de feiten bekend waren is er gekeken hoe deze ingedeeld moesten worden. De hele context van feiten wordt beschreven met behulp van dimensies. Dimensies zijn alle gegevens waar geen berekening op kan worden uitgevoerd maar die gebruikt worden om de feiten in categorieën in te delen zodat ze inzichtelijk worden gepresenteerd. Elke dimensie beschrijft een invalshoek van een feit. Er is dus gekeken welke dimensies relevant zijn. Alle niet relevante dimensies zijn niet opgenomen. Dimensies waren relevant als deze voorkwamen in de informatiebehoefte van de proceseigenaren. Het vinden van dimensies vereist een bepaalde denkwijze. De dimensies zijn gevonden door te redeneren hoe de proceseigenaren hun informatie wilden zien en door gebruik te maken van een dimensietabel. In deze dimensietabel staan alle meetwaarden en dimensies onder



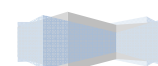
elkaar in de verticale as. In de horizontale as zijn de managementvragen (informatiebehoefte) te vinden. Op deze manier kan er met kruisjes worden aangegeven welke managementvraag betrekking heeft op welke feiten en welke dimensies. Op deze manier ontstond een overzicht van managementvragen, feiten en dimensies. Omdat er was besloten niet alle informatiebehoefte uit te werken maar enkel degene met de hoogste prioriteit is er een feitenmodel ontstaan met de vijf eerder genoemde zaken en de bijhorende dimensies:

Feit	Dimensies
Forecast directe kosten	Tijd, businessunit
Forecast indirecte kosten	Tijd, businessunit
Forecast Bruto toegevoegde waarde	Tijd, businessunit
Budget directe kosten	Tijd, businessunit
Budget indirecte kosten	Tijd, businessunit
Budget Bruto toegevoegde waarde	Tijd, businessunit
Forecast Ebit	Tijd, businessunit
Budget Ebit	Tijd, businessunit
Verbrand budget	Project
Projectprijs	Project
Omzet Directe Medewerkers	Tijd
Orderportefeuille	Tijd

Normaal gesproken is er een bepaalde hiërarchie aanwezig binnen een dimensie, bijvoorbeeld bij de tijdsdimensie. Een hiërarchie maakt het mogelijk in te zoomen op een bepaald feit. Zo kan er bijvoorbeeld bij het feit “verkoop” in een supermarkt worden gekeken naar de verkopen per jaar en kan er worden ingezoomd naar de verkopen per week. Hiërarchie was bij dit project niet van toepassing omdat de proceseigenaren alleen behoefte hadden aan maandcijfers. Het totaal voor het betreffende jaar zou automatisch worden weergegeven zodra alle maanden werden weergegeven. Hiërarchie binnen de businessunit-dimensie was ook niet van toepassing.

Bij het bepalen van de feiten is er gewerkt volgens de top-down methode. Dit betekent dat er is gekeken welke behoefte er is en daarna is er pas gekeken of dit te realiseren is met de bestaande data. Het voordeel hiervan is dat er gekeken wordt wat de organisatie echt wil, terwijl er bij de bottom-up benadering in het brongegevensmodel wordt gekeken welke dimensies relevant lijken en wordt er geen rekening gehouden met de informatiebehoefte. Zo is er tijdens de interviews met de proceseigenaren gevraagd in welke vorm zij hun rapportage of KPI wilden hebben. Zo kon er de relatie gelegd worden tussen feiten en dimensies. Hoe ze de data wilden zien was namelijk een dimensie. Zo werd duidelijk dat de financiële cijfers (directe-, indirecte kosten, bruto toegevoegde waarde, EBIT) getoond moesten worden in maanden, maar dat men ook behoefte had aan een weergave per businessunit, twee dimensies dus. Al had men aangegeven de data in kwartalen te willen zien had dit ook gerealiseerd kunnen worden. Daarnaast was voor de projecten behoefte aan inzicht in de budgetverbranding. De enige dimensie die hier van toepassing was: project. Er wordt namelijk gekeken naar verbruikt budget per project. Als laatste was er nog de orderportefeuille. Er was behoefte aan een overzicht tussen de orderkansen, gevallen orders en de orders die nodig waren (omzet directe medewerkers). Dit wilde men ook graag per maand zien. Een logische dimensie was in dit geval dus ook “tijd”.

De laatste stap voor dit hoofdstuk was om te kijken welke gegevens relevant waren om op te slaan binnen de dimensies. Met andere woorden: welke gegevens nemen we mee en welke laten we



vallen. Was er bijvoorbeeld de dimensie “medewerker”, dan moest er voor deze dimensie bepaald worden welke data relevant was om mee te nemen, bijvoorbeeld: achternaam, geslacht, leeftijd. Zo moest er voor iedere dimensie bepaald worden welke attributen meegenomen moesten worden. Voor de tijdsdimensie was enkel het attribuut “maand” relevant. Jaartal was niet relevant omdat we slechts de data uit het huidige jaar moesten presenteren. Overige data die met een tijd te maken had zoals: uren, werkdagen, kwartalen, etc. namen we ook niet mee.

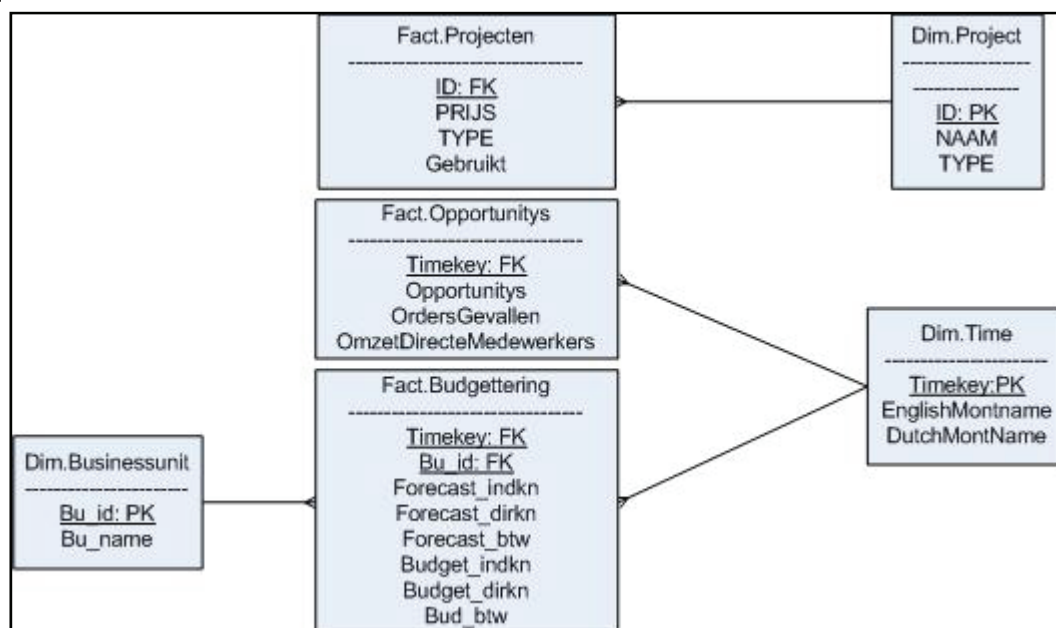
Voor de projectdimensie waren er een aantal relevante attributen. Ten eerste was de naam van het project belangrijk. Ook de prijs van het project moest worden meegenomen om het gebruikte budget af te kunnen zetten tegen de prijs. Tot slot was voor de businessunitdimensie slechts de naam van de businessunit interessant om mee te nemen. Men kan in de toekomst eventueel medewerkers gaan koppelen aan een businessunit om verder in te kunnen zoomen op bepaalde gegevens. Dit was echter in eerste instantie niet noodzakelijk, daarom is ervoor gekozen om enkel gegevens mee te nemen die ook gewenst waren door de opdrachtgever.

Nu alle feiten en dimensies bekend waren kon er gestart worden met het opstellen van een sterschema, het echte ontwerp van het datawarehouse.

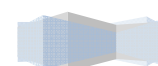
7.5 Opstellen sterschema

Het volgende ontwerp dat terug te vinden is in het designrapport is die voor het datawarehouse. Er is gekozen voor de ontwerptechniek van Kimball, het sterschema. Dit is het meest gebruikte model om een datawarehouse te ontwerpen. Het sterschema is eerst op papier ontworpen. Per feit of feitengroep zou een tabel binnen het datawarehouse ontstaan, zo ook voor dit sterschema. Hieronder is het sterschema te vinden van het datawarehouse van ETTU. Onder het sterschema is uitgelegd waarom er is gekozen voor de tabellen zoals deze zijn opgesteld.

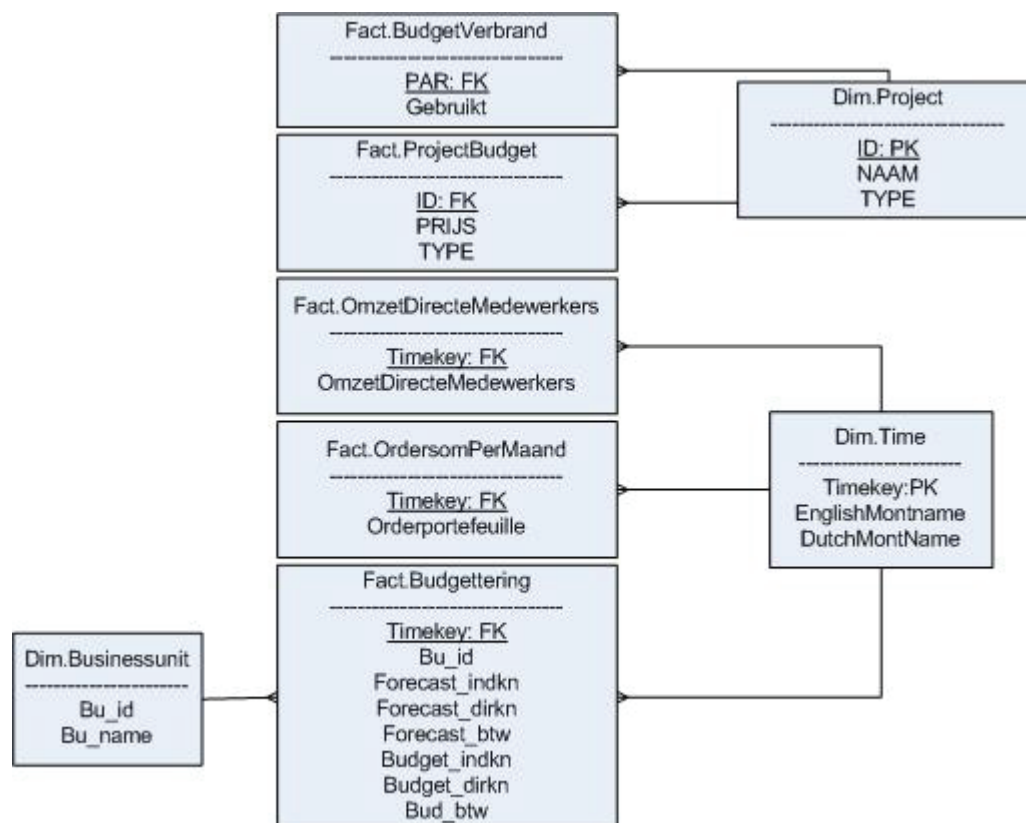
Uiteindelijk is ervoor gekozen het ontwerp iets aan te passen zodat de tabellen eenvoudiger gevuld konden worden bij het overzetten van de data. Het model hieronder is het model dat in eerste instantie ontworpen was, daaronder is het model te vinden zodat deze uiteindelijk is opgesteld voor ETTU.



Figuur 20 Sterschema datawarehouse



Het verschil met het eerste en het tweede model is dat het tweede model meer tabellen bevat. Bij het ontwikkelen van de oplossing liep ik diverse keren tegen het probleem aan dat de data niet goed werd overgezet of werd overschrijven. Om niet teveel tijd te verliezen heb ik er daarom voor gekozen de tabellen te splitsen.



Figuur 21 Sterschema datawarehouse uiteindelijke versie

Uitleg sterschema

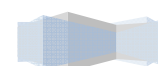
In de meeste gevallen bestaat een sterschema uit maximaal twee feitentabellen en omliggende dimensies. Ik heb er uiteindelijk voor gekozen alle meetgroepen in een aparte feitentabel te plaatsen omdat dit makkelijker was het overzetten van gegevens tijdens het ETL-proces. Alle feitentabellen zijn voorzien van het voorvoegsel “fact” en de dimensietabellen van “Dim”. Op deze manier is de begrijpbaarheid van het datawarehouse vergroot. Onderstaande drie tabellen geven een uitleg over de tabellen. Deze drie tabellen zijn ingedeeld naar onderwerp. Zo bevat de eerste tabel gegevens over de financiële zaken, de tweede over de orders en de laatste bevat informatie over projecten.

Fact.budgettering

Deze tabel bevat alle financiële gegevens die uit het budgetteringssysteem afkomstig zijn. Dit zijn de forecast- en budgetgegevens. Omdat er besloten was deze gegevens in te delen in maanden en in businessunit moesten er verwijzingen worden opgenomen naar deze dimensies door middel van een foreignkey Timekey en een foreignkey Bu_id.

Fact.Opportunities

Deze tabel bevat de het bedrag aan orderkansen in een bepaalde tijdseenheid, in dit geval per maand.



Opportunities

De som van alle orders in een maand die mogelijk gaan vallen. Dit is de kans * orderbedrag. De maand van vallen is de sluitdatum.

OmzetDirecteMedewerkers

Omdat de orderportefeuille afgezet moest worden tegen de omzet van de medewerkers moest hiervoor ook een feitentabel worden opgesteld. Deze is gekoppeld met de tijdsdimensie.

OrdersGevallen

De opdrachtgever wilde graag een overzicht van alle orderkansen, gevallen orders en omzet directe medewerkers voor een bepaalde maand, vandaar dat deze feitentabel is opgenomen in het datawarehouse. Het bevat gegevens over de gevallen orders per maand. Tevens is deze gekoppeld met de tijdsdimensie voor een weergave per maand.

Fact.Projecten

De prijzen van de projecten moesten ergens opgeslagen worden. Er is gekozen om dit in een feitentabel op te slaan. Deze tabel is gekoppeld met de projectdimensie. Als sleutel is het projectID in deze tabel opgenomen.

Id

Het ID van het project om deze te kunnen koppelen met de projectdimensie.

Prijs

De prijs van een project zoals deze is vastgelegd in het urensysteem

Type

Controlelevel, wordt verder niet gebruikt in het project

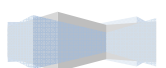
Gebruikt

In deze kolom staat de geaggregeerde data van het gebruikte budget van een project. Alvorens deze tabel gevuld wordt zullen er enkele berekening moeten worden uitgevoerd.

Het bovenstaande ontwerp is niet zomaar uit de losse hand ontstaan. Er is nagedacht over hoe de data het beste kon worden opgeslagen zodat het ook begrijpbaar is voor de medewerkers van ETTU. Als er in een later stadium iets aangepast moet worden dan moet dit niet al teveel moeite gedaan kunnen worden, ook voor mensen die niet erg veel verstand hebben van de tools. Ook is er rekening gehouden met het feit dat er geen overbodige gegevens meegenomen worden in het datawarehouse om de performance hoog te houden. Tijdens het ontwerpen is er ook het een en ander getest door SQL-queries uit te voeren en de uitkomsten te vergelijken met de gegevens in de bronsystemen.

7.6 Voorbereidende activiteiten ETL-schema's

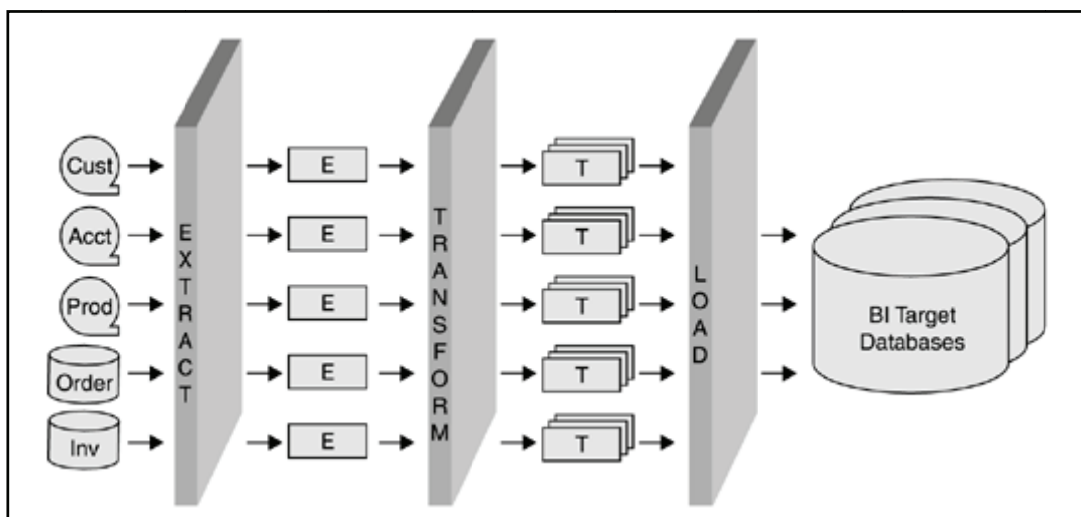
Om de juiste rapportages te kunnen vervaardigen is data nodig die in diverse databronnen is opgeslagen. De benodigde data kwam uit drie bronnen, vandaar dat er het een en andere overgezet moest worden richting het datawarehouse. Dit is gedaan met behulp van een ETL-proces (extract, transform en load). Dit betekent dat de data uit de huidige bron wordt gehaald (extract), wordt



omgezet indien dat nodig was (transform) en in het datawarehouse wordt geplaatst (load). Het softwarepakket dat hiervoor gebruikt is: Microsoft SQL Server 2005, onderdeel Integration Services (SSIS). Dit pakket was standaard aanwezig binnen ETTU zodat er geen toolselectie heeft plaatsgevonden. Hieronder is een afbeelding te vinden ter verduidelijking van de werking van het ETL-proces. Aan de linkerkant staan de diverse bronsystemen, aan de rechterkant het datawarehouse. Daartussen zijn de drie bewerkingen terug te vinden. Het ETL-ontwerp bestaat uit een aantal te volgen stappen:

- ETL-proces voorbereiden;
- ETL-schema's opstellen;
- ETL tools testen;
- ETL-schema beschrijvingen;
- ETL-package bouwen.

Stap 1, 2 en 3 zijn door elkaar uitgevoerd. Tijdens het ontwerp is er met de tools geëxperimenteerd om de mogelijkheden te bekijken en is er tevens bekeken welke data omgezet moest worden in het voorbereidende proces. Wellicht waren er voldoende mogelijkheden binnen SSIS om de ideeën te realiseren. Toch was het handig om te weten wat wel en niet mogelijk was, mede omdat ik nog geen ervaring had met de tool.

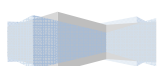


Figuur 22 Werking ETL-proces

Een ETL-proces wordt vooraf gegaan door de data te herformatteren, redundante gegevens te herkennen en data op te schonen. Daarna moet er een keuze worden gemaakt welke methode wordt toegepast om de data te laden. Dit is ook toegepast voor dit project.

Data herformatteren

Omdat het budgetteringssysteem in Excel was opgesteld stonden de gegevens verspreid over de tabbladen. Omdat SSIS de data op deze manier niet kon uitlezen moest de data hergeschikt worden naar een leesbaar formaat zodat het overgezet kon worden naar het datawarehouse. Er is daarom een nieuw tabblad ontworpen binnen het budgetsysteem waarin de gegevens in een bruikbaar formaat zijn vastgelegd. In dit nieuwe tabblad zijn verwijzingsvelden opgenomen naar de gegevens uit de andere tabbladen. Verandert er iets aan de velden in het originele tabblad als gevolg van een wijziging, dan veranderen de gegevens in het nieuwe tabblad automatisch mee. Dit nieuwe tabblad is op onzichtbaar gezet zodat de gebruiker niet doorheeft dat dit tabblad aanwezig is. Hij heeft dit tabblad namelijk niet nodig voor zijn werk. Daarnaast is er een tweede tabblad opgenomen met de



omzet directe medewerkers-gegevens per maand. Dit is gedaan omdat de gegevens over de omzet van de medewerkers ook overgezet moesten worden naar het datawarehouse.

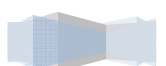
Onderstaande twee tabellen tonen welke kolommen zijn opgenomen in de nieuwe tabbladen. De eerste rij van het tabblad zijn kolomnamen. Later kan er tijdens het ETL-proces worden aangegeven dat de eerste rij geen gegevens bevat maar kolomnamen zodat de namen niet worden uitgelezen en worden overgezet.

Tabblad: Warehouse
Maand: integer
Businessunit: integer
Forecast indirecte kosten: money
Forecast directe kosten: money
BTW: money
Gebudgetteerde BTW: money
Gebudgetteerde Indirecte kosten: money
Gebudgetteerde directe kosten: money
EBit Forecast: money
EBit budget: money

Tabblad: Warehouse-2
Maandnummer: integer
OmzetDirecteMedewerkers:money

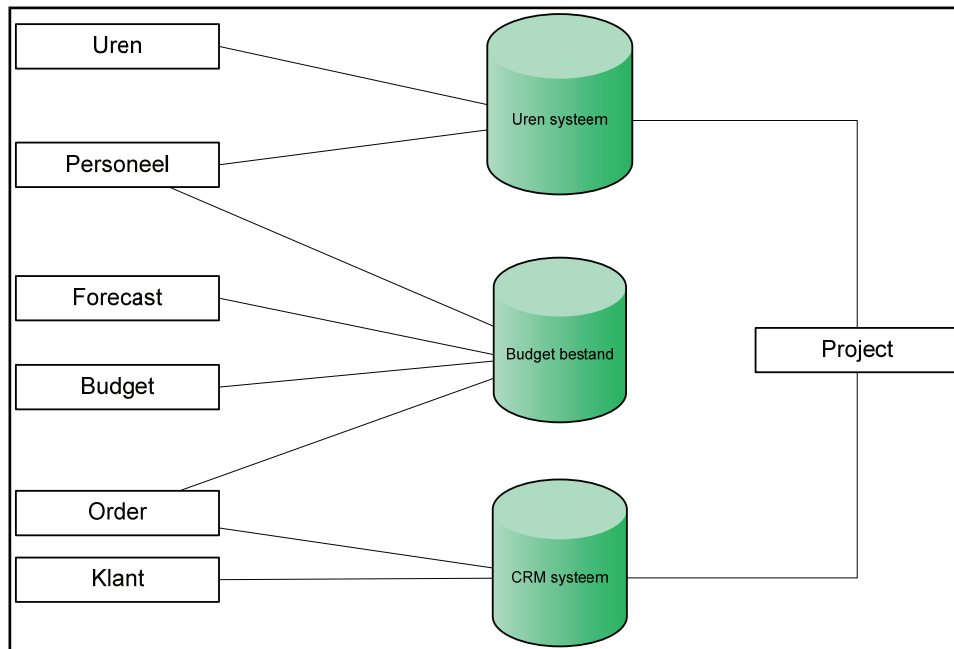
Herformatteren van de data uit het urensysteem was eenvoudiger. Omdat deze data in een MYSQL database stond kon er worden volstaan met het overzetten naar een tijdelijke tabel in het datawarehouse om van daaruit de juiste bewerkingen te kunnen toepassen. SQL Server biedt namelijk meer mogelijkheden en commando's dan Mysql.

De laatste brondatabase was vastgelegd in SQL server zodat herformatteren niet nodig was.



Redundante gegevens herkennen

Omdat het belangrijk is om geen dubbele gegevens mee te nemen in het datawarehouse waardoor gegevens inconsistent kunnen worden was het belangrijk hier naar te kijken. Zoals in het bronnenmodel is te zien is waren er enkele overlappende gegevens aanwezig tussen de diverse bronsystemen. Onderstaand model toont het bronnenmodel:



Figuur 23 Bronnenmodel ETTU

Enkele overlappende gegevensvelden waren:

- Order;
- Project;
- Personeel.

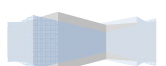
Tijdens het uitvoeren van de data-analyse bleek dat er geen gegevens gebruikt zouden worden die in twee of meer systemen voorkwamen. In dit geval zouden ordergegevens redundant kunnen zijn wanneer deze in het datawarehouse geladen moesten worden vanuit twee bronsystemen en waarbij gegevens in beide systemen een andere betekenis zouden hebben. Ordergegevens zouden enkel uit het CRM-systeem worden gehaald, personeelsgegevens uit het urensysteem en projectinformatie ook uit het urensysteem. Dat betekende dat er in de bronsystemen wel redundante gegevens voorkwamen maar dat deze niet in het datawarehouse zouden worden overgezet.

Data opschonen

De daaropvolgende voorbereidende stap was het opschonen van bevuilde data. Aangezien er geen vuile data gebruikt werd was deze stap overbodig. Tijdens de data-analyse is bepaald dat er een paar projecten in het urensysteem voorkwamen die dubbel waren ingevoerd. Ik heb dit aangegeven en men zou daar in het vervolg rekening mee houden.

Data laden

Het laden van data is hetzelfde als de data uit de bronsystemen halen of uitlezen. Er zijn drie mogelijkheden die gebruikt kunnen worden bij het laden van data: initial load, historical load en incremental load. Er moest in deze fase van het project bekeken worden welke methode het beste kon worden toegepast.



Initial load laadt alle data uit een bronsysteem, inclusief de data die gebruikt wordt (operationele data). Historical load is hetzelfde als een initial load, alleen wordt er ook historische archiefdata meegenomen. Tot slot is er de incremental load. Deze methode zorgt ervoor dat alleen nieuwe data wordt overgezet. Er wordt dus gekeken welke data er al aanwezig is in het datawarehouse en vervolgens wordt alleen de data geladen die nog niet in het datawarehouse aanwezig is.

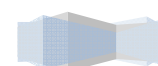
Er is uiteindelijk gekozen voor methode een, een initial load. Ten eerste bleek dit de eenvoudigste methode te zijn. Er hoefde geen vergelijkingen te worden uitgevoerd tussen bron- en doeltabel om te zien welke data nieuw was. Dit is wel het geval bij een incremental load. Daarnaast vergde een initial load weinig performance van het systeem omdat het iedere keer om een kleine hoeveelheid data ging. Zo hoefde er uit het budgetsysteem slechts 50 regels geladen te worden, het CRM-systeem 300 regels en het urensysteem 4000 regels. Als er bedacht wordt dat bij grote systemen soms wel miljoenen regels geladen worden is dit een mini-laadproces. Historical load was geen optie omdat er geen historische gegevens aanwezig waren (statische gegevens op bijv. tapes).

7.7 ETL-schema budgetsysteem

Het budgetsysteem was gemaakt in een Excelbestand. De benodigde gegevens waren inmiddels gestructureerd opgeslagen door het uitvoeren van de vorige stap. Nu was het zaak om te kijken hoe deze data overgezet kon worden naar het eerder ontworpen datawarehouse. Het was verstandig om tijdens het ontwerp direct wat toolverkenning uit te voeren om te kijken wat de mogelijkheden waren. SSIS is een erg uitgebreide applicatie met veel mogelijkheden. Er zijn enkele tutorials uitgevoerd om de basisfunctionaliteiten van het programma te leren kennen. Zo is er gebruikt gemaakt van wat voorbeelden van Microsoft om te zien hoe een ETL-package was gemaakt. Op deze manier raakte ik al snel vertrouwd met de applicatie.

Een ETL-package (output van SSIS) is een executable bestandje die in Windows gedraaid kan worden. Dit bestandje bevat alle instructies die meegegeven worden bij het ontwerpen van een ETL-workflow. De eerste stap bij het maken van ETL-package was een connectie aanmaken met een databron, in dit geval Excel. Het budgetsysteem was een office 2007 bestand met de extensie .xlsx. Dit was ook de oorzaak van een nieuw probleem. SSIS heeft een connectionmanager die ervoor zorgt dat een connectie kan worden gelegd met een bepaalde bron. Zo zijn er diverse connectionmanagers aanwezig voor Oracle, SQL, Excel, ODBC, etc. Logischerwijs viel de keuze op Excel omdat het om een Excelbestand ging. Bij het connecten kon er echter alleen een Excelbestand gekozen worden die in de oude versie was gemaakt. De nieuwe versie van Microsoft Office 2007 werd dus niet ondersteund door SSIS. Omdat ik me niet kon voorstellen dat dit niet mogelijk was ben ik gaan zoeken op internet. Ik kwam daar al snel dezelfde vraag tegen die ik mijzelf stelde en ook de oplossing. Door het upgraden van SQL Server 2005 naar Service Pack 2 zou er een nieuwe connectionmanager aan de lijst worden toegevoegd die de gebruiker in staat moest stellen een connectie te leggen met Office 2007 bestanden. Na het maken van een back-up van het systeem is SP2 geïnstalleerd en de nieuwe connectionmanager was inderdaad aanwezig in de lijst. Omdat niet bekend was hoe deze precies te gebruiken was is het internet weer geraadpleegd. Daar is een voorbeeld ETL-package gevonden waarin een connectie was gemaakt met een office 2007 bestand. De instellingen daaruit zijn overgenomen wat resulteerde in een werkende connectie.

ETTU maakt gebruik van Microsoft Sharepoint 2007 portal (intranet omgeving) voor de beheersing van documenten. Ook het budgetsysteem was opgeslagen in dit systeem. Bij het leggen van een eerdere connectie met dit bestand was het bestand opgeslagen op de locale schijf van de server. Opnieuw dook er een probleem op. Het was een raadsel hoe connectie kon worden gemaakt met

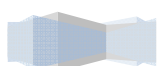


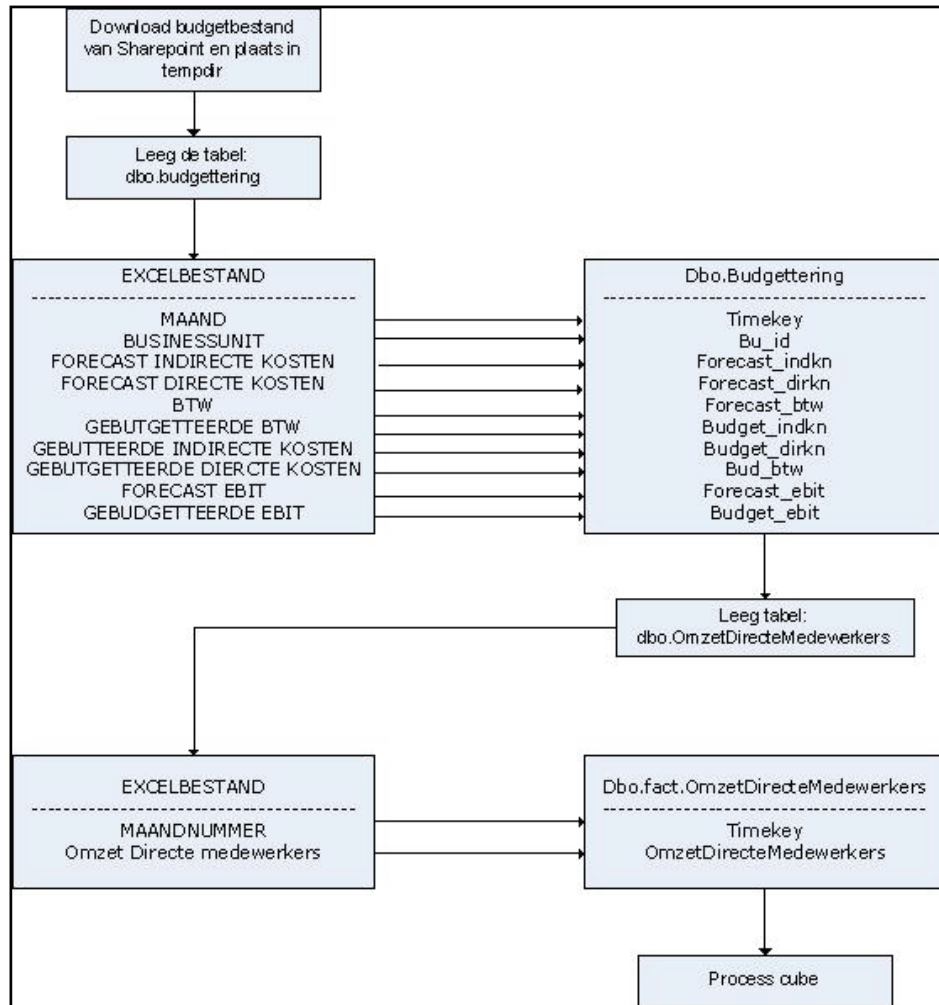
een bestand dat was opgeslagen in Microsoft Sharepoint. Om hierachter te komen is er weer gezocht op internet. Op een forum was een artikel te lezen over iemand met hetzelfde probleem. Hij had daarvoor een script ontwikkeld die te koppelen was aan een ETL-package. Dit script zorgde ervoor dat aan het begin van het ETL-package het bestand werd gedownload naar een map op de lokale schijf. Daarna kon er met de Office 2007 connectionmanager verbinding worden gemaakt met het bestand. Een goede en eenvoudige oplossing. Het script bleek na het uitvoeren van een test goed te functioneren. Deze test omvatte het downloaden van een bestand naar de lokale schijf van de server. De ontwikkelaar was in staat eenvoudig een locatie op te geven voor het bestand (http locatie) en de map waarin het bestand gezet moest worden kon ook eenvoudig worden aangepast. Nu kampte ik met de vraag of het een probleem was als het financiële bestand in een tijdelijke map opgeslagen was op de server. Deze vraag is voorgelegd aan een proceseigenaar. Hij heeft aangegeven dat dit geen probleem was omdat er verder niemand op de server kon komen.

Daarnaast is er bij het ontwerp voor gekozen om de doeltabellen in het datawarehouse eerst te legen alvorens te vullen met nieuwe data. Een hoofdreden waarom dit gedaan is:

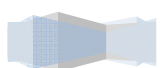
- **Zeker weten dat er geen dubbele gegevens in de doeltabellen staan;**

Verder kostte dit proces weinig performance omdat het datawarehouse weinig data bevat (en ooit zal bevatten) en dit een eenvoudige methode is om zeker te weten dat er geen dubbele gegevens in het datawarehouse komen te staan. Bij het testen van de ETL-tool liep ik iedere keer tegen het probleem op dat hij de data netjes achter de reeds aanwezige data plaatste. Alle gegevens kwamen op deze manier dubbel voor in het systeem. Dit bleek een snelle en effectieve oplossing aangezien er toch gebruik gemaakt werd van een initial load waarbij alle data wordt overgezet. Onderstaand schema toont het ETL-proces voor het budgetsysteem. Voor meer uitleg over de stappen wordt doorverwezen naar bijlage D.



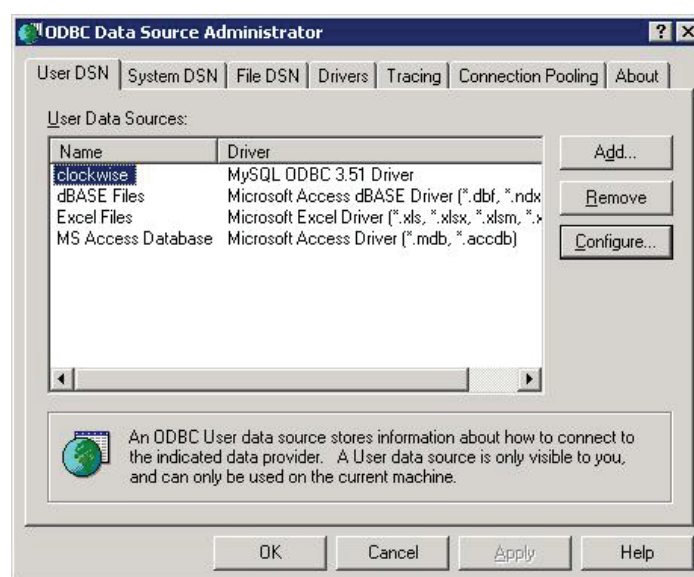


Figuur 24 ETL-schema budgetstelsysteem



7.8 ETL-schema urensysteem

Het volgende ETL-schema was die van het urensysteem. Het urensysteem was een MYSQL database die op een server draaide. Ook hier is tijdens het ontwerpen direct gekeken welke mogelijkheden de tool bood. Als eerste moest er worden begonnen met een connectie maken. SSIS bevatte een ODBC connectionmanager waarmee connectie kon worden gelegd met een database. Ik wist niet hoe dit werkte dus ben gaan zoeken op het internet. Om te kunnen connecten met een ODBC (Open Database Connectivity) database moest er een zogenaamde ODBC-driver worden geïnstalleerd op de windows server. In deze ODBC-driver kon een verbinding worden gelegd met de MYSQL database.



Figuur 25 ODBC driver

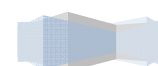
Deze verbinding kon vervolgens worden gebruikt in SSIS. Deze verbinding is aangemaakt en getest. Nu de verbinding er was kon ik SQL-queries uitvoeren op de MYSQL-database van het urensysteem. Ik zal niet alle activiteiten toelichten die zijn uitgevoerd. Voor uitleg over het ETL-schema verwijs ik door naar bijlage D. Daar is ook het complete ETL-schema te vinden.

7.9 ETL-schema CRM-systeem

Het CRM-systeem was het enige systeem dat gebruik maakte van een Microsoft SQL database. Onderstaande zaken zouden gemeten gaan worden:

- Orderkansen huidige jaar per maand;
- Gevallen orders huidige jaar per maand.

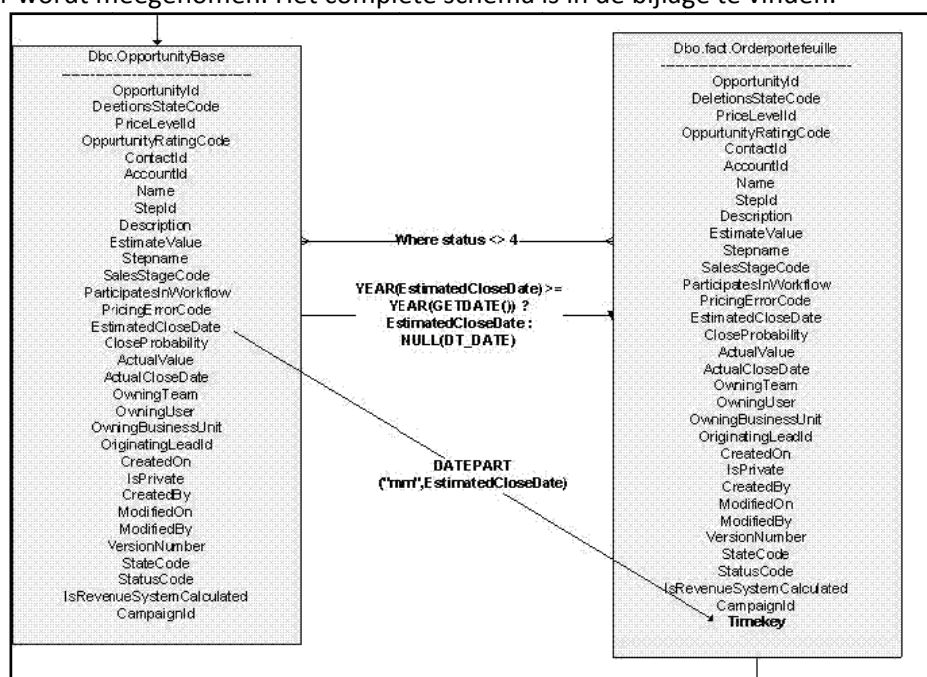
Zoals te zien is werd hier ook gebruik gemaakt van de tijdsdimensie maand. Uiteraard was er geen sleutel aanwezig in de brontabel zodat de nieuwe tabel in het datawarehouse gekoppeld kon worden met de tijdsdimensie. Er moest dus op de een of andere manier een sleutel worden toegevoegd. De oplossing was een surrogaatsleutel gebruiken voor het maandnummer. Dit is een unieke waarde die aan iedere rij van de feitentabel wordt toegevoegd. Deze sleutel wordt de primaire sleutel van de dimensietabel tijd en wordt gebruikt om een relatie te leggen tussen de feitentabel en de dimensietabel. Voordelen: de sleutel beveiligd het systeem tegen veranderingen in de bronsystemen, maakt integratie mogelijk van data uit meerdere bronnen (als deze bijvoorbeeld verschillende



sleutels gebruiken), en een integer als surrogaatsleutel zorgt voor een hogere performance in analysis services (compact, klein formaat feitentabel). Dit kon worden verzorgd door SSIS met behulp van de “derive column” workflowactie . Dit commando zorgt ervoor dat er een rij wordt toegevoegd in de doeltabel waarbij een voorwaarde wordt meegegeven. Ik had zelf bedacht om het maandnummer uit het datumveld te halen van de sluitdatum van de order. Het maandnummer kon dan gekoppeld worden aan de tijdsdimensie. Een collega programmeur heeft op dat moment hulp geboden bij het invoeren van de statements. Het resultaat was dat er een nieuwe rij kon worden toegevoegd aan de doeltabel met het maandnummer als timekey zodat een koppeling mogelijk werd.

Vervolgens moest er ook rekening worden gehouden met het feit dat we alleen de orders wilden zien uit het huidige jaar, oude data uit bijvoorbeeld 2006 was niet relevant. Er is voor gekozen om alle data over te zetten in een tijdelijke tabel, vervolgens de timekey toe te voegen en tot slot een voorwaarde mee te geven die let op het jaartal uit het veld “EstimatedCloseDate”. Daarbij kijkt het programma naar het jaartal op het moment van uitvoeren en vergelijkt dit met de waarde uit het veld van EstimatedCloseDate. Als het jaartal niet overeenkomt worden deze velden gevuld met een NULL-waarde en vervolgens verwijderd. Resultaat: een tabel met alle orders uit het huidige jaartal die voorzien is van een timekey. Nog een puntje waar rekening mee is gehouden is het feit dat alleen orders meegenomen moesten worden die niet verloren waren (die niet doorgingen). Dit was te controleren met behulp van het statusveld. Alle orders die geen status=4 hadden werden overgezet.

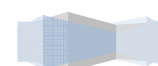
Onderstaand figuur toont een stukje van het ETL-schema van het CRM-systeem naar het datawarehouse. Hierin is te zien hoe de timekey wordt toegevoegd en hoe alleen de data uit het huidige jaar wordt meegenomen. Het complete schema is in de bijlage te vinden.



Figuur 26 Timekey toevoegen CRM

De tijdelijke tabel die nu ontstond bevatte:

- Orders uit het huidige jaar;
- Orders die een kans maken;
- Orders die zijn gevallen;
- Een surrogaatsleutel naar de tijdsdimensie.



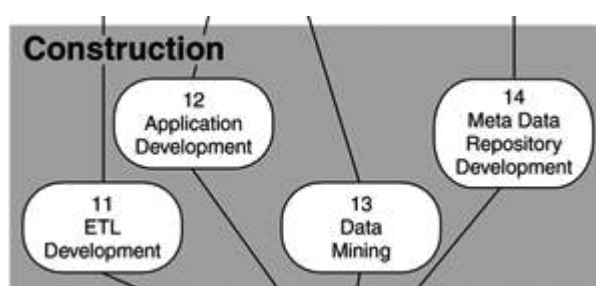
8.0 Fase 5: Construction

8.1 Inleiding

Het uiteindelijke product dat gerealiseerd zou worden voor ETTU en met name de opdrachtgever was een werkende oplossing. Een werkende oplossing hield in: een managementdashboard in Microsoft Sharepoint 2007 die up-to-date moest zijn en dus automatisch gevuld moest worden met de juiste data. Daarbij moesten de achterliggende ETL-processen goed ingericht zijn zodat het datawarehouse automatisch ge-update zou worden.

Een werkende oplossing kon alleen gerealiseerd worden wanneer alle schema's goed waren ontworpen en er gedegen kennis in huis was om de oplossing te ontwikkelen. Omdat er nog weinig kennis aanwezig was van de ontwikkeltools moest deze kennis eerst verkregen worden. Tijdens het ontwikkelen van de eerder opgestelde schema's is er al geëxperimenteerd met SSIS. In deze fase van het project moest er ook kennis worden opgedaan van de tool om het datawarehouse in te ontwikkelen, Microsoft SQL Server Analysis Services (SSAS).

Onderstaande afbeelding toont de vier deelfasen van de Construction fase. Deelfasen 13 en 14 zijn niet van toepassing op dit project omdat er geen data-mining gebruikt wordt en er geen meta data repository is gebruikt in verband met de kleine hoeveelheid gegevens.

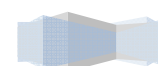


Figuur 27 Fase 5: Construction

8.2 Literatuurstudie en toolverkenning SQL Server 2005

Omdat al voor het afstuderen bekend was dat SQL Server 2005 gebruikt zou gaan worden heb ik direct een boek aangeschaft, namelijk: The Microsoft Data Warehouse Toolkit van de Kimball Group. Dit boek leert hoe SQL Server 2005 te gebruiken is. Het gaat niet diep in op alle ontwerp- en ontwikkeldetails van SQL Server maar vertelt welke onderdelen voor welke doeleinden te gebruiken zijn. Verder legt het uit hoe een sterschema ontworpen kan worden en waar op gelet moet worden bij het ontwerpen van een sterschema of bijvoorbeeld een ETL-proces. Alvorens te starten met de toolverkenning zijn de eerste hoofdstukken van het boek bestudeerd. De volgende onderwerpen stonden hierin behandeld:

- Waar moet je op letten bij het bepalen van de business requirements?
- Hoe maak je een sterschema?
- Uit welke tools bestaat SQL Server 2005?
- Hoe ontwikkel je een datawarehouse?
- Hoe ontwerp je een ETL-proces?
- Hoe ontwikkel je een ETL-proces?
- Hoe maak je een OLAP-database in Analysis Services?



In een vroeg stadium van het project is er al gestart met het bestuderen van dit boek. Er ontstond meer kennis van Business Intelligence en de mogelijkheden van de applicatie. SQL Server 2005 bestond uit een aantal onderdelen. Naast de databaseserver bestond deze applicatie uit de volgende drie onderdelen:

- Analysis Services (SSAS)
- Integration Services (SSIS)
- Reporting Services (SSRS)

De eerste was geschikt voor het ontwerpen van een datawarehouse, de tweede voor het ontwikkelen van ETL-processen en de laatste voor het maken van rapportages. Wellicht zouden alle drie deze onderdelen gebruikt worden voor het project. Daarom is er besloten alle drie de tools te verkennen op hun mogelijkheden.

Database-verkenning

Omdat SQL Server 2005 erg veelzijdig en uitgebreid is was het lastig om een start te maken, waar moest ik beginnen? Het leek verstandig om te beginnen wat ervaring op te doen met de database zelf. Dit was immers de fundering van het project. SQL Server bevatte een Management Studio om een connectie te maken met een databaseserver. Vervolgens werd een overzicht getoond van alle aanwezige databases op de server, beveiligingsinstellingen en noem maar op.

Door rond te klikken in de applicatie en te experimenteren werd ik al snel vertrouwd met de werking ervan. De belangrijkste zaken die ik onder de knie moest krijgen voor het project:

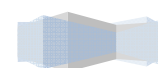
- Hoe maak je een database en tabellen aan?
- Hoe zet je data in deze tabellen?
- Hoe voer je een query uit op de database?
- Hoe creëer je een view?

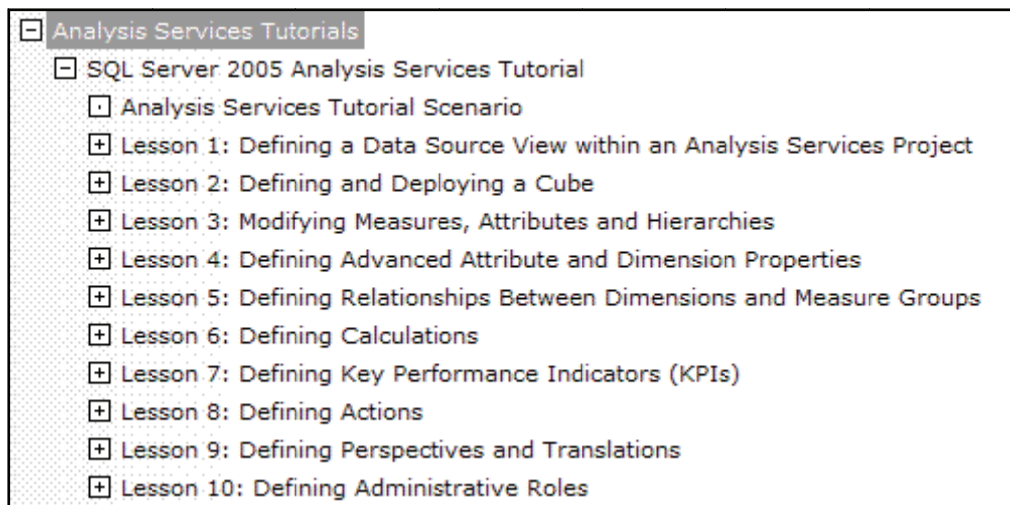
Bovenstaande vier activiteiten zijn uitgevoerd voor de database die inmiddels op de testserver was geïnstalleerd, de Adventureworks database van Microsoft.

SSAS-verkenning

Nadat ik bekend was met de werking van de database ben ik gaan kijken hoe de tool werkte waarin het datawarehouse gebouwd zou gaan worden, Analysis Services. De eerste keer dat dit programma werd opgestart wist ik niet waar ik moest beginnen, alles was nieuw en ik wist niet hoe het werkte. Hoe kon ik mijn ontwerp gaan bouwen in deze tool? Om het snelst de werking van deze tool te leren leek het verstandig een tutorial (lessen) te volgen. Aangezien er voldoende tijd was ingebouwd voor de constructionfase paste dit prima in het project. Daarnaast biedt Microsoft een aantal presentaties aan op haar website in de vorm van Windows Mediaplayer Movies en een aantal Virtual Labs.

Microsoft heeft op haar MSDN pagina een zeer uitgebreide tutorials voor Analysis services. Hiervoor was testdata benodigd van het welbekende Adventureworks voorbeeld. Deze data was al geïnstalleerd op de testserver waardoor ik direct aan de slag kon. Ik vond het belangrijk om (bijna) de hele tutorial te volgen omdat ik nog niet bekend was met de applicatie. Ook wilde ik graag zoveel mogelijk functies van het programma leren en weten wat het programma allemaal kon. De tutorial zou mij hier een stuk verder in kunnen helpen. De gehele tutorial bestond uit tien lessen die ieder een bepaald onderwerp behandelden. Onderstaand figuur toont de opbouw van de tutorial.





Figuur 28 Analysis Services Tutorial

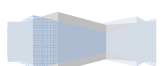
Zoals te zien is in bovenstaand figuur behandelt het tutorial alle belangrijke zaken die iemand moet weten als hij een datawarehouse wil maken. Er wordt ingegaan op het maken van een cube, relaties leggen, meetgroepen aanmaken, KPI's definiëren, etc. Het volgen van deze tutorial en het maken van de bijbehorende oefeningen heeft veel tijd in beslag genomen. Het heeft er wel voor gezorgd dat ik vervolgens beter in staat was een datawarehouse te bouwen. Het was namelijk geen doen om "zomaar wat" te proberen in deze applicatie omdat er gigantisch veel mogelijkheden en functionaliteiten aanwezig waren.

SSIS-verkenning

Vervolgens is er gekeken hoe het ETL-programma werkte, SQL Server Integration Services (SSIS). De toolverkenning van SSIS heeft plaatsgevonden tijdens het ontwerpen van de ETL-schema's. Ik heb niet al aan het begin van het project een toolverkenning gedaan zoals bij de andere onderdelen. De reden hiervoor was simpel. Ik heb bij het testen van SSAS en SSRS gebruik gemaakt van gegevens die handmatig waren vastgelegd in de database. Het was dus niet noodzakelijk om aan het begin van het project al gegevens over te zetten met behulp van een ETL-proces.

SSRS-verkenning

Omdat het nog onduidelijk was of deze tool gebruikt zou gaan worden leek het mij toch handig hier even naar te kijken. Omdat een proceseigenaar Gauge charts wilde zien in zijn rapportage en deze niet aanwezig waren binnen SSRS ben ik gaan zoeken op het internet. Daar kwam ik op de website van Dundas.com een addon tegen voor SSRS. De demoversie hiervan is geïnstalleerd en er zijn rapporten gemaakt met deze chart. De gauge chart hoefde uiteindelijk niet gebruikt te worden omdat deze 2000 dollar kostte. Dit is besloten bij de presentatie van de voorlopige versie.



8.3 Toolverkenning Microsoft Sharepoint 2007

Microsoft Sharepoint 2007 was de portal waarin het dashboard gemaakt moest worden, de front-end van de oplossing die getoond moest worden aan de eindgebruikers. Om te begrijpen hoe deze applicatie werkte heb ik rondgekeken en wat zaken getest zoals het maken van een verbinding met een cube, een rapport uit SSRS en een excelbestand. Zo was er een omgeving aanwezig die “reports” heette. Hierin konden rapporten en dashboardpagina’s worden aangemaakt. Er was ook al een testdashboard aanwezig die een gevoel gaf hoe het in elkaar stak. De volgende onderdelen waren in het testdashboard te vinden:

- Connectie met een Excel 2007 bestand;
- Gedefinieerde KPI’s met lampjes;
- Gegevens uit Excel 2007 en een grafiek.

Het was uiteindelijk de bedoeling om de data te presenteren in sharepoint vanuit een Excel werkboek. Dit voorbeeld gaf een goede indruk van de werking. Het voorbeeld maakte gebruik van statische data uit Excel. Excel had dus geen connectie met een datawarehouse, hetgeen bij dit project wel het geval zou zijn. Ik zou dus zelf moeten uitzoeken hoe ik de data vanuit de cube kon weergeven in Excel en uiteindelijk ook wel stukjes van de excelsheet.

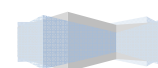
Om te kijken hoe een connectie kon worden gelegd tussen sharepoint en een externe bron heb ik wat testen uitgevoerd. Sharepoint bestaat uit pagina’s waarin webpart worden weergegeven. Een webpart is een stukje functionaliteit, vergelijkbaar met een component of module. Zo was er een webpart aanwezig waarmee KPI’s getoond konden worden. Ik heb een connectie gemaakt met het adventureworks datawarehouse (demo datawarehouse). De keuze viel op de connectie “OLAP CUBE” die standaard kon worden gekozen. Door wat te testen kreeg ik de werking snel onder de knie. De connectie kon worden gemaakt en de KPI’s die in de CUBE waren gedefinieerd konden worden weergegeven in sharepoint. Vervolgens heb ik wat filters aangemaakt om de waarden te kunnen filteren. Dit zou later eventueel gebruikt gaan worden om te filteren op businessunit of project.

8.4 Constructie ETL-packages

Nu de tools verkend waren en de ontwerpen klaarlagen was het tijd voor de ontwikkeling van de oplossing. Het leek mij verstandig om eerst de ETL-packages te maken alvorens het datawarehouse te bouwen. De data moest namelijk eerst in het systeem staan voordat er rapportages gemaakt konden worden. Tijdens het testen heb ik gewerkt met dummie-data maar nu was het tijd om de data uit de live-omgeving te gaan gebruiken.

Bij het bouwen van de ETL-packages heb ik me gehouden aan het opgestelde ontwerp. De meeste problemen waren al tijdens de verkenningsfase getackeld. Het bouwen ging dus redelijk vlot omdat er al flink was geëxperimenteerd. Bij de constructie van de ETL-packages heb ik ervoor gekozen de tabellen in het datawarehouse automatisch aan te laten maken. Bij het ontwikkelen van een ETL-packages moest er een bronsysteem en een doelsysteem worden geselecteerd. Bij het selecteren van het doelsysteem kon daarvoor gekozen worden (create_new_table). Het voordeel hiervan was dat ik niet eerst tabellen hoefde te definiëren in SQL Server maar dit automatisch gebeurde zodat ook direct de veldtypen goed stonden.

Op deze manier zijn alle drie de ETL-packages ontworpen. Per bronsysteem is er een ETL-package gemaakt zodat deze afzonderlijk gedraaid konden worden. Het ene systeem hoeft namelijk minder vaak geladen te worden dan het andere.



8.5 Constructie datawarehouse

Nadat alle drie de ETL-packages waren ontwikkeld en ook op de juiste manier functioneerden was het tijd om het datawarehouse te bouwen in SSAS. De SQL-database waarin de data was opgeslagen was inmiddels gevuld door de ETL-packages. Doordat ik alle tutorials had gevolgd was de omgeving van SSAS vertrouwd. Ik ben gestart met het maken van een connectie met de database. Daarna kon er via een wizard een cube worden gedefinieerd. Omdat het ontwerp van de database goed was kon de wizard automatisch de sleutels en relaties herkennen. Ook herkende hij automatisch de feiten.

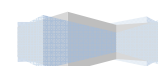
De volgende stap was het definiëren van dimensies. Ook hier was een wizard voor aanwezig. De wizard leidde me door de stappen heen. Er moest een tabel worden geselecteerd waarin de dimensie te vinden was. Vervolgens werden de relaties automatisch gelegd zodra de namen van de sleutels overeenkwamen. Als dit niet het geval was kon er in een view handmatig een relatie worden gedefinieerd. Het handige was dat er in de SQL-database geen relaties aanwezig hoefden te zijn. Deze kunnen namelijk handmatig worden ingesteld in de cube.

Nu de feiten en dimensies gebouwd waren heb ik enkele testen uitgevoerd om te kijken of ik de juiste gegevens boven tafel kon krijgen. SSAS bevat een browserscherm waarin resultaten kunnen worden getoond. Zo kon ik een feit naar het midden van het scherm slepen en vervolgens een dimensie. De data werd nu op de juiste manier gepresenteerd. In onderstaand voorbeeld worden er twee meetwaarden weergegeven in de tijdsdimensie. Links zien we de maanden van het huidige jaar, daarnaast de som van de opportuniteiten en tot slot vinden we in de laatste kolom het bedrag aan gevallen orders.

Drop Filter Fields Here		Drop Column Fields Here	
Dutch Mont Name ▼	Orderportefeuille - Fact Order Som Per Maand	Orderportefeuille	
Januari		€ 56.660,00	
Februari	€ 2.000,00	€ 211.500,00	
Maart		€ 80.500,00	
April	€ 17.000,00	€ 18.000,00	
Mei	€ 138.500,00	€ 27.906,00	
Juni	€ 149.000,00		
Juli	€ 108.750,00		
Augustus	€ 46.700,00		
September	€ 68.750,00		
Oktober	€ 4.000,00		
December	€ 10.000,00		
Grand Total	€ 544.700,00	€ 394.566,00	

Figuur 29 meetwaarden en dimensie tijd

Sommige meetwaarden moesten in twee dimensies getoond worden, bijvoorbeeld de financiële gegevens. Deze moest per businessunit worden getoond en per maand. Door het definiëren van de juiste dimensies met de juiste feiten kon onderstaande draaitabel worden weergegeven.



Drop Filter Fields Here		Drop Column Fields Here
Dutch Mont Name	Bu Name	FC Directe kosten als percentage VD BTW
Januari	Applications	33,58%
	Configurations	24,20%
	Detachering	0,00%
	Total	31,26%
Februari		36,11%
Maart		40,39%
April		49,05%
Mei		44,66%
Juni		43,88%
Juli		50,77%
Augustus		48,36%
September		53,14%
Oktober		40,39%
November		41,60%
December		54,95%
Grand Total		44,49%

Figuur 30 meetwaarden en dimensies tijd en businessunit

Calculations

Een belangrijke functie binnen SSAS was de calculation functie. Met deze functie konden berekening worden gedefinieerd die vervolgens als feit te gebruiken waren bij het weergeven van data. De gegevens met betrekking tot directe- en indirecte kosten waren in de database opgeslagen als getallen. Uiteindelijk moesten deze gegevens worden weergegeven in percentages zoals in bovenstaand figuur is te zien. Dit kon worden gerealiseerd door gebruik te maken van de calculation functionaliteit.

In de calculation functie heb ik een aantal berekening gedefinieerd. De indirecte kosten moesten bijvoorbeeld als percentage van de bruto toegevoegde waarde worden weergegeven. Dit kon worden gerealiseerd door de volgende formule:

“ ([Measures].[Forecast Dirkn] / [Measures].[Forecast Btw]) “

We delen hier de forecast directe kosten door de forecast BTW. Daarnaast definiëren we dat de uitkomst van het type `percent` moet zijn. Op deze manier zijn er vier berekening gemaakt, namelijk:

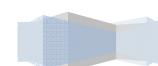
- Forecast directe kosten;
- Forecast indirecte kosten;
- Gebudgetteerde directe kosten;
- Gebudgetteerde indirecte kosten.

Key Performance Indicators

Binnen SSAS was het mogelijk KPI's te definiëren. Er kon een doelwaarde worden ingesteld en een waarde waarbinnen het lampje op groen blijft branden. Daarnaast kon er een formule worden ingesteld die zorgt dat de trendpijl omhoog of naar beneden ging staan. In eerste instantie heb ik alle KPI's binnen SSAS gedefinieerd. Dit bleek te werken. Bij het testen liep ik tegen de volgende problemen op:

- Lampje kon alleen op rood of op groen staan, oranje leek niet mogelijk;
- Eurotekens werden dollars;
- Er kon alleen gebruik worden gemaakt van lampjes, niet van grafieken.

Het probleem met de dollar is opgelost door een collega programmeur in te schakelen die het probleem heeft opgelost door gebruik te maken van een bepaalde parameter in de KPI die met een landencode te maken had. De andere twee waren goede redenen waarom ik er uiteindelijk voor gekozen heb de KPI's niet in SSAS te definiëren maar door gebruik te maken van de business Intelligence functies in de nieuwste versie van Excel.



8.6 Constructie Dashboard in Microsoft Sharepoint 2007

Voor het tonen van de data uit het datawarehouse binnen een Sharepoint 2007 omgeving kon er gebruik worden gemaakt van verschillende methoden. Zouden we gebruik gaan maken van een KPI-webpart, van reporting services webparts of van Excel Services?

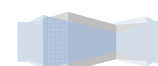
De keuze is uiteindelijk gevallen op Excel services. Excel services is een webpart binnen sharepoint 2007 die het mogelijk maakt stukjes data uit een excelsheet weer te geven in sharepoint. Deze methode bood alle functies die nodig waren en ETTU genoot de voorkeur voor deze methode. ETTU wilde namelijk experimenteren met nieuwe functies in office 2007. Nog meer redenen waarom er niet is gekozen voor de andere twee methoden:

- KPI webpart biedt geen mogelijkheden voor charts;
- KPI webpart maakt gebruik van KPI's uit SSAS dus enkel rood en groen;
- Reporting services is lastig aan te passen door ETTU;
- KPI's definiëren in reporting service is lastig en vergt kennis van programmeren;
- Rapporten uit Reporting services zijn statisch.

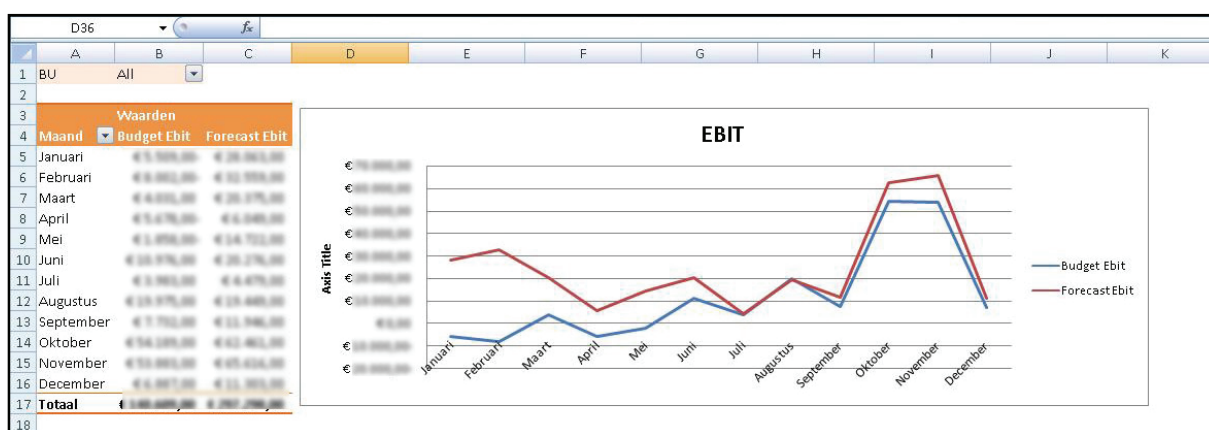
Excel 2007 is voor iedereen goed te begrijpen zodat een kleine aanpassing eenvoudig is te realiseren, bijvoorbeeld de kleur van een grafiek of de filters op de draaitabellen. In Excel 2007 zijn er een aantal tabbladen aangemaakt, voor iedere gewenste rapportage een tabblad. Daarna moest er een verbinding komen tussen de cube en Excel. Omdat ik niet wist hoe ik dit moest realiseren heb ik op google gezocht. Daar kwam ik de oplossing al snel tegen. Er moest een bestand worden aangemaakt, een odc file, die een connectie moest maken met de cube. Excel herkende daarna automatisch alle feiten, calculations, KPI's en dimensies die in de cube waren vastgelegd. Vervolgens konden de feiten en gedefinieerde berekening vanuit de cube worden geselecteerd, samen met de juiste dimensies. Op deze manier ontstond er een draaitabel. Er konden filters worden ingesteld en er kon een grafiek worden gekoppeld aan de data voor een grafische weergave. Nog een mooie optie was het definiëren van KPI's. Er zijn KPI's gedefinieerd in de kostensheets. In alle velden waarbij het kostenpercentage hoger was dan 40% moest ook rood lampje branden en onder de 37% groen. Daartussen moest het lampje oranje worden. In eerste instantie lukte het niet om de waarden te definiëren. Omdat het om percentages ging moest er bijvoorbeeld niet de waarde '40' worden ingesteld maar '0,4' voor een rood lampje. Dat was iets waar je net op moet komen.

B28		fx	
	A	B	C
1	BU	All	
2			
3	Waarden		
4	Maand	Budget Indirecte Kosten	Forecast Indirecte Kosten
5	Januari	57,81%	31,81%
6	Februari	58,09%	35,62%
7	Maart	45,99%	43,16%
8	April	53,57%	45,42%
9	Mei	52,03%	43,76%
10	Juni	65,57%	40,45%
11	Juli	45,55%	45,23%
12	Augustus	38,71%	38,25%
13	September	38,91%	38,33%
14	Oktober	28,54%	27,84%
15	November	26,69%	25,97%
16	December	38,52%	38,03%
17	Totaal	43,05%	36,75%

Figuur 31 Excel 2007 draaitabel met KPI's

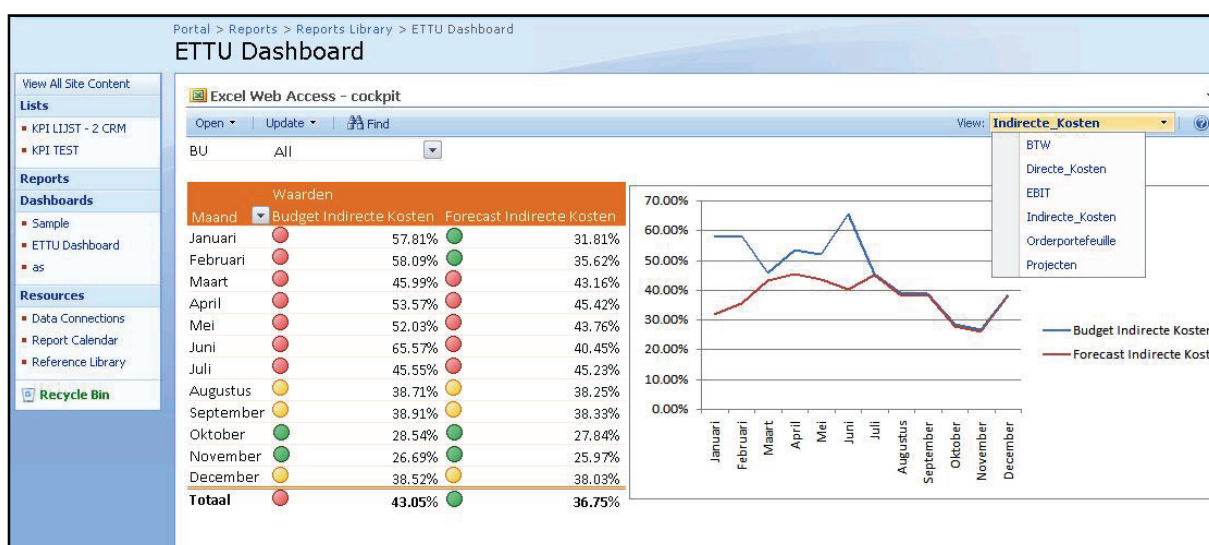


De volgende stap was om de juiste data uit het Excelbestand weer te geven in sharepoint. Niet alle inhoud van de excel tabbladen moest namelijk getoond worden maar enkel dat stuk dat gegevens bevatte en waar de chart te vinden was. Binnen Excel kon er een bereik worden geselecteerd. Aan dit bereik kon een zogeheten 'named range' worden gekoppeld, een bepaalde naam. Op deze manier is er voor ieder tabblad een named range aangemaakt met het juiste bereik. Onderstaand voorbeeld toont een draaitabel binnen Excel in combinatie met een chart en een filter voor de businessunit. Daarnaast kunnen er eventueel maanden worden weggelaten door de maandfilter te gebruiken. De data is afkomstig uit SSAS en wordt iedere keer herladen zodra het bestand geopend wordt.

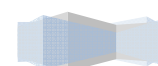


Figuur 32 Excel 2007 grafiek

Tot slot moest het Excelbestand gekoppeld/opgeslagen worden aan/in sharepoint. Dit kon door middel van de publicatiefunctie in Excel. Daarbij kon er worden aangegeven welke named ranges meegenomen moesten worden. Hier zijn de ranges gekozen die eerder zijn gedefinieerd. Door het bestand te publiceren werd het opgeslagen binnen sharepoint. Na de publicatie van het bestand binnen sharepoint moest het nog getoond worden in sharepoint door middel van het Excel services webpart. Deze moest geactiveerd worden op de dashboardpagina. Vervolgens moest de locatie van het Excelbestand worden opgegeven en de standaard named calculation die getoond moest worden bij het openen van de pagina. Het mooie was dat er automatisch een slider werd getoond met alle named ranges uit Excel. De gebruiker kon dus zelf kiezen welke rapportage hij wilde zien. Onderstaande afbeelding toont het resultaat met rechtsboven de slider.



Figuur 33 Dashboard in Sharepoint 2007

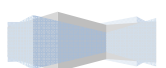


Ik vond het erg belangrijk om de data duidelijk en overzichtelijk te presenteren. Men moest snel kunnen bepalen wat de stand van zaken was. Daarom heb ik ervoor gekozen om aan de ene kant gebruik te maken van draaitabellen en aan de andere kant van grafieken. Wat waren de voordelen hiervan?

- Grafieken tonen snel de stand van zaken maar zijn niet nauwkeurig;
- Getallen geven iets nauwkeurig weer ;
- Getallen kunnen in combinatie met KPI's worden gebruikt, in dit geval lampjes.

Een aantal rapportages moesten de forecast en het budget tegen elkaar afzetten. Deze verhoudingen konden perfect worden weergegeven door middel van bovenstaande grafiek. Een staafdiagram kon dit niet zo mooi als een lijndiagram. Nemen we het gebruikte budget per project is een staafdiagram een betere oplossing omdat deze de twee staafjes naast elkaar zet, links het budget en rechts het verbruikte budget. Zo is in een oogopslag te zien hoe een project ervoor staat. De orderportefeuille was tot slot het beste weer te geven met een staafdiagram. Een staaf met de opportuniteiten, staaf twee toon de gevallen orders en staaf drie de omzet medewerkers. Ik heb er ook voor gekozen de grafieken zo eenvoudig mogelijk te houden. Maximaal drie duidelijk te onderscheiden kleuren en geen 3d-grafieken, dat zou alleen bedriegend kunnen werken .

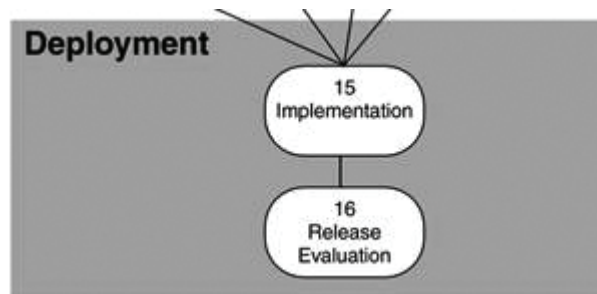
Voor de KPI's was de wens om lampjes te maken. Daarom heb ik drie lampjes gekozen; rood, oranje, groen. De proceseigenaar van deze KPI heeft de streefwaarden voor de kleuren bepaald zodat deze ingevoerd konden worden in Excel 2007. Deze lampjes kunnen ook eenvoudig worden vervangen door andere symbolen omdat ze in Excel zijn gedefinieerd.



9.0 Fase 6: Deployment

9.1 Inleiding

De laatste fase van het project was de deploymentfase waarin het project overgezet zou gaan worden in de live omgeving. Daarnaast vindt er een release evaluation plaats waarin er wordt teruggekeken op het project. Deze laatste was niet van toepassing op dit project. In deze fase is het systeem getest en is er een implementation report geschreven voor de opdrachtgever.



Figuur 34 Fase 6: deployment

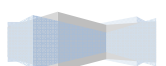
9.2 Testen

Omdat een juiste gegevensuitvoer cruciaal is moest het systeem getest worden. Dit is gedaan door de uitvoer te vergelijken met data uit de bronsystemen.

Als eerste is de uitvoer van de financiële data getest. Dit was eenvoudig omdat de data uit het budgetsysteem vergeleken kon worden met de uitvoer. Kwamen deze cijfers overeen dan klopte de uitvoer.

De volgende test was de controle op de uitvoer van de ordergegevens. Om deze uitvoer te testen is er een query op het bronsysteem uitgevoerd waarbij alle orders van het huidige jaar moesten worden weergegeven. Dit waren alle orders met een status ongelijk aan 'lost order'. Vervolgens heb ik voor vier maanden de orderportefeuille uitgerekend. Voor de opportuniteiten was dit: "Bedrag * Slagingspercentage" bij elkaar opgeteld voor een bepaalde maand in het jaar 2007. Als dit bedrag overeenkwam met de uitvoer van het systeem dan klopte het. Voor de gevallen orders ging het hetzelfde, alleen hoefde er nu geen vermenigvuldiging plaats te vinden met een slagingspercentage, de kans was immers altijd 100 bij een 'won order'. Hier was ook een foutje geslopen in het systeem. Ik had namelijk voor de orderportefeuille voor gevallen orders ook een vermenigvuldiging ingesteld maal de slagingskans. Deze fout zal in het ETL-package van CRM. Door het uitvoeren van deze test is deze grote fout er uitgehaald.

Tot slot was er nog de projecteninformatie met het gebruikte budget per project. De uitvoer hiervan is getest door een query uit te voeren die de uren vermenigvuldigde met het uurtarief en dat voor een bepaald project. De uitvoer moest overeenkomen met deze query. Dit bleek te kloppen.



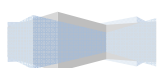
9.3 Implementatie

De implementatie van het systeem hoefde niet door mij te worden verzorgd. Om de technische medewerker van ETTU op weg te helpen had ik besloten een implementation report voor hem te maken. Dit is geen milestone van een bepaalde fase in het project maar is ontstaan uit eigen initiatief. Het implementation report moest de medewerker op weg helpen het systeem te implementeren in de live-omgeving.

In het implementation report zijn de volgende onderwerpen behandeld:

- Opgeleverde projectbestanden en hun werking;
- Werking van de SQL-database;
- Omschrijving van de ETL-packages en variabelen;
- Aanpassingen die nodig zijn aan het budgetsysteem;
- Werking van de cube;
- Werking van Excel in combinatie met sharepoint.

Met behulp van dit document zou men in staat moeten zijn het project live te zetten. Ik vond het een groot risico om het project zonder mijn hulp te implementeren. Ik heb diverse malen aangeboden mee te helpen met de implementatie. Omdat men geen tijd had heb ik in week 16 voor de laatste keer aangeboden te helpen. Omdat men erg druk was met projecten was dit de laatste kans voor hen om het met mijn hulp te doen. Daarna zou men zelf de verantwoordelijkheid moeten nemen voor de implementatie.



10.0 Evaluatie

10.1 Inleiding

In de voorgaande hoofdstukken heb ik geprobeerd een goede beschrijving te geven van mijn werkzaamheden tijdens het afstuderen. Deze hoofdstukken zijn gericht op de verschillende activiteiten die mijn afstuderen vormde.

In dit hoofdstuk heb ik het geheel aan eindproducten, het proces en de planning geëvalueerd.

10.2 Evaluatie eindproducten

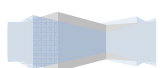
Het eindproduct dat men zal gaan gebruiken bij de firma ETTU is de cockpit die geïntegreerd is binnen hun Sharepoint 2007 portaal. Daarbij was het van belang dat de meest recente gegevens getoond werden. Het vullen van het datawarehouse moest volledig automatisch gebeuren.

Het eerste opgeleverde product was het Business Requirements Report waarin alle informatiebehoefte terug te vinden is. Ik denk dat het rapport een goede weergave is van de informatiebehoefte van ETTU. Door aan het begin van het project overleg gevoerd te hebben met de diverse belanghebbenden en door het begrijpen van de werkwijze van ETTU ontstond een beeld van de informatiebehoefte. Dit is beschreven en gevisualiseerd in het rapport. Het voordeel van dit rapport was dat de opdrachtgever en ik overeenstemming hadden over de spelende informatiebehoefte. Ik moest begrijpen waar men naar toe wilde. Daarnaast kon ik met dit rapport de opdrachtgever laten zien hoe zijn informatiebehoefte kon worden vertaald in een rapportage of grafiek.

De volgende milestone, het Business Design Report, was van essentieel belang voor een juiste ontwikkeling van de oplossing. In dit rapport zijn alle modellen van de oplossing terug te vinden. Dit document wordt gebruikt om het systeem te bouwen. In dit geval heb ik het systeem ook zelf gebouwd maar het document zou gebruikt kunnen worden door een derde persoon die het systeem zou moeten bouwen. Daarnaast geeft het document voldoende informatie over het systeem. Het zou bijvoorbeeld gebruikt kunnen worden wanneer er een aanpassing moet worden gedaan.

Het systeem dat uiteindelijk is opgeleverd voldoet aan alle eisen die vooraf zijn gedefinieerd. In het Business Requirements Report zijn alle weer te geven onderdelen uitgewerkt en is er overeenstemming bereikt over deze rapportages. Deze zijn uiteindelijk allemaal terug te vinden in de oplossing. Voordelen voor ETTU bij het gebruik van de oplossing:

- Beter zicht op de prospectieve situatie;
- Projectvolgning mogelijk;
- Sneller kunnen ingrijpen.



Die zaken waarop men wilde sturen zijn allemaal geïntegreerd in de business Intelligence oplossing. Enkele nadelen van de oplossing:

- Beperkte mogelijkheden tot een drill-down;
- Geen historische gegevens;
- Geen interne kennis aanwezig tot aanpassingen.

Het eerste nadeel zie ik niet als een groot nadeel. ETTU had behoefte aan een kort en bondig overzicht van data. Daarbij kan men drillen op van een totaaloverzicht naar een maandoverzicht en van alle businessunits tot aan een bepaalde businessunit. Voor sommige rapportages is het niet nuttig om te drillen. Het zou bijvoorbeeld niet nuttig zijn om het verbruikte budget voor een project te drillen. De gegevens die dan tevoorschijn zouden komen zijn de geschreven uren van de medewerkers. Wat wel interessant had kunnen zijn is om de orderportefeuille te drillen of de kostengegevens. Al zou ik het project over moeten doen zou ik mij meer gericht hebben op deze drill-down functie. Ik heb bij het ontwikkelen echter rekening gehouden met de wensen vanuit de business en deze wens was niet naar voren gekomen. Uiteraard is het de taak van iemand als mij om verder te denken richting de toekomst. Ik heb hier wel aangedacht alleen was het project al lastig genoeg om uit te voeren.

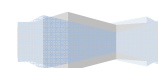
Historische gegevens waren niet aanwezig binnen ETTU omdat het bedrijf nog geen jaar bestond. Het systeem is er echter ook niet op gebouwd om historische gegevens op te slaan. Ook dit heeft te maken met de informatiebehoefte. Men had behoefte aan een overzicht van het huidige jaar waarbij historische gegevens niet van belang waren. Wellicht kan dit in de toekomst wel belangrijk worden, bijvoorbeeld wanneer men data-mining wil gaan toepassen. Bij data-mining gaat het systeem voorspelling maken richting de toekomst aan de hand van historische gegevens. Dit viel echter buiten de scope van het project.

Tot slot was er het implementation report. Dit rapport moest de technische medewerker van ETTU op weg helpen de oplossing te implementeren binnen de live-omgeving. Het rapport geeft voldoende informatie om de oplossing te implementeren zonder hulp van mij. Toch kan ik met dit rapport niet garanderen dat de oplossing geïmplementeerd kan worden. Er kunnen namelijk bepaalde dingen niet instaan die cruciaal zouden kunnen zijn voor de implementatie en die voor mij erg logisch zijn. Het was echter de keuze aan ETTU om zelfstandig een implementatie uit te voeren.

10.3 Evaluatie proces

Om dit project succesvol te kunnen voltooien is er gebruik gemaakt van een structurele aanpak. Het project is opgesplitst in deelfasen waarbij er is geprobeerd de activiteiten uit te voeren die van toepassing waren op het project. Niet iedere activiteit die beschreven stond was namelijk van toepassing omdat het om een relatief klein project ging waarin weinig data ontsloten werd uit de bronsystemen. De gehanteerde methode gaat uit van een gemiddeld BI-project waarin duizenden of zelfs miljoenen records worden overgezet naar een datawarehouse.

Ook is er tijdens de ontwikkeling van de business Intelligence oplossing zoveel mogelijk rekening gehouden met de eisen en wensen vanuit de business. Aan het begin van het project is er regelmatig contact geweest met de opdrachtgever en de proceseigenaren om de informatiebehoefte zo goed mogelijk te kunnen bepalen. Dit was ook de belangrijkste fase uit het project en vormde als het ware de fundering van het project. De informatiebehoefte die zich vertaalde in een set van KPI's en rapportages is gepresenteerd aan de proceseigenaren om overeenstemming te bereiken over hetgeen dat weergegeven moest worden in de cockpit. Zodra deze informatiebehoefte duidelijk was



en iedereen het eens was over de KPI's en rapportages kon er pas echt gestart worden met het project.

Iedere milestone die opgeleverd moest worden is teruggekoppeld aan de opdrachtgever. Het plan van aanpak was de eerste milestone die besproken is met de opdrachtgever. Dit plan gaf duidelijkheid aan beide partijen en kon worden gezien als een contract. Daarna is het rapport met de informatiebehoefte opgeleverd en besproken zodat de opdrachtgever ook wist welke ideeën ik had en hoe ik de rapportages wilde vormgeven. In dit rapport stond hoe de informatiebehoefte zich vertaalde in managementrapportages en KPI's. De opdrachtgever had op dat moment kunnen aangeven als hij andere ideeën had over de rapportages of KPI's. Opmerkingen en nieuwe ideeën bleven op dat moment uit. Het gevolg hiervan was dat de opdrachtgever in de laatste fase van het project een dimensie erbij wilde in de rapportage die de orderportefeuille weergaf. Ik heb aangegeven te kijken of dit nog in het project paste. Uiteindelijk heb ik besloten dit niet meer te realiseren omdat alle ontwerpen en het project zelf al klaar waren en er anders grote aanpassingen nodig waren. Dit paste ook niet meer binnen het tijdsbestek van het project. Misschien had ik de opdrachtgever meer bewust moeten maken van het feit dat kleine wijzigingen grote technische aanpassingen nodig hadden in het systeem. Waarschijnlijk heeft hij dit onderschat waardoor hij op het laatste moment nog met een aanpassing kwam.

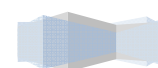
Wanneer men (op tijd) aangaf een aanpassing te willen in het systeem heb ik hier rekening mee gehouden. Zo wilde men nog een rapportage in de cockpit toen ik de voorlopige versie had laten zien, hetgeen een flinke aanpassing nodig had in het systeem. Dit waren de EBIT gegevens die in eerste instantie niet in het business design report stonden. Ik denk dat ik dit rapport uiteindelijk ook met alle proceseigenaren had moeten bespreken en niet alleen met de opdrachtgever. Gelukkig was er op dat moment nog voldoende tijd en ruimte om veranderingen door te voeren.

Wanneer ik dit project nog een keer zou moeten uitvoeren zou ik het waarschijnlijk op dezelfde manier aanpakken. Belangrijke zaken waar ik meer op zou letten is de business meer bewust maken van impact van een kleine verandering op het systeem. Daarnaast zou ik de belanghebbenden meer bewust maken van het belang van ondersteuning bij de implementatie. Ik heb gemerkt dat men erg makkelijk dacht over het invoeren van het systeem. Dit bleek ook uit het feit dat men in de laatste week van het project nog steeds geen tijd had voor invoering, ook toen ik mijn hulp had aangeboden. Wellicht zal de implementatie hierdoor niet vlekkeloos verlopen. Bewustwording of wellicht het prikken van een definitieve datum had dit kunnen voorkomen.

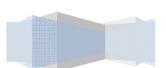
10.4 Evaluatie planning

In de beginfase van het project is er een globale planning opgesteld waarin de zes fasen terug te vinden zijn. Per fase is er een begindatum en een einddatum vastgelegd. Daarnaast is er een detailplanning opgesteld waarin alle activiteiten per fase terug te vinden zijn. Ik had mijzelf voorgenomen me in grote lijnen te houden aan de planning. Uit ervaring weet ik dat ik niet een echte planner ben. Toch was ik mij bewust van de noodzaak van een planning, al was het maar als houvast. De planning moest mij helpen het project op tijd af te krijgen.

Alle producten zijn binnen de gestelde afstudeerperiode afgerond. Doordat ik erg nieuwsgierig was naar de mogelijkheden van SQL Server 2005 ben ik zelfs buiten werktijd met deze tool bezig geweest. Daardoor had ik de tool snel onder de knie en kon ik eerder beginnen met het bouwen van de oplossing. Hierdoor had ik aan het einde van het project wat meer ruimte om de rapportages fraaier te maken. Ook had ik hierdoor meer tijd om mij bezig te houden met de mogelijkheden van



bijvoorbeeld reporting services wat uiteindelijk niet gebruikt zal worden. Kennis opdoen is echter nooit weggegooid tijd.



Bijlage A: Opdrachtomschrijving

Inleiding (organisatorische omgeving, kader, historie)

ETTU is een organisatie van ongeveer een jaar oud. Er zijn op dit moment 20 mensen werkzaam en het is de bedoeling dat ETTU binnen een aantal jaren zal groeien tot 50 man. ETTU is opgebouwd uit twee businessunits. De ene businessunit houdt zich bezig met de ontwikkeling van mobiele toepassingen. Een concreet voorbeeld is een applicatie voor digitale werkbonden. Zodra de monteur klaar is met zijn werkzaamheden voert hij in zijn PDA is welke onderdelen hij heeft gebruikt en hoe lang hij bezig is geweest. Dit wordt automatisch via GPRS naar de back-office gestuurd. Vervolgens krijgt hij digitaal zijn volgende opdracht aangeleverd en wordt automatisch zijn Tomtom navigatie ingesteld.

De tweede businessunit ontwikkelt en beheert Microsoft Sharepoint portals voor bedrijven. Deze portals worden gebruikt om projectdocumenten beschikbaar te maken voor projectmedewerkers. Dit zijn flexibele portals die zich kunnen aanpassen aan de eindgebruiker.

Probleemstelling

ETTU wordt gestuurd met behulp van een budget dat begin van het jaar wordt vastgesteld. In dit budget staan de opbrengsten en kosten van het aankomende jaar. Daarnaast maakt men gebruik van een forecast (voorspelling) waarin de verwachtingen voor een bepaalde periode worden gezet. Hierin zet men bijvoorbeeld hoeveel mensen er aangenomen worden en welke projecten doorgaan. Om de juiste interne beslissingen te kunnen nemen en voorspellingen te maken richting toekomst is informatie nodig die op dit moment niet inzichtelijk is en alleen kan worden verkregen door data uit verschillende systemen met elkaar te vergelijken. Op dit moment zijn de volgende zaken niet inzichtelijk:

- Hoeveel omzet is er gebudgetteerd en hoeveel geld in orders is er in de portefeuille aanwezig voor een bepaalde maand?
- Hoeveel budget is er daadwerkelijk (op een bepaald moment) gebruikt voor een bepaald project?
- Het verschil tussen bepaalde kostenposten/opbrengsten en datgene dat is gebudgetteerd.
- Overall probleemstelling: doen we het goed ten opzichte van wat is gebudgetteerd?

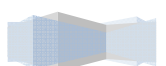
Als deze informatie inzichtelijk is kan men de organisatie intern beter sturen. In situatie 1 kan men meer orders proberen binnen te halen als blijkt dat er te weinig orders zijn. In geval twee projecten beter plannen of meer tijd reserveren als blijkt dat projecten altijd buiten budget treden. In geval drie kosten beter beheersen.

Nog een bijkomstigheid is dat Business Intelligence een nieuw onderdeel is binnen Sharepoint 2007. ETTU wil graag meer weten over de mogelijkheden van dit nieuwe onderdeel. Op dit moment is er nog geen kennis aanwezig over Business Intelligence in combinatie met Sharepoint 2007. Men wil dit in de toekomst aan klanten aanbieden als blijkt dat het intern goed werkt.

Doelstelling van de opdracht

Dit project kent twee doelstellingen.

- Een interne Business Intelligence oplossing ontwikkelen zodat informatie overzichtelijk gepresenteerd wordt in een managementdashboard. Hierdoor kunnen proceseigenaren betere beslissingen nemen over het aankomende jaar.



- De organisatie kennis laten opdoen van de mogelijkheden van Business Intelligence binnen Sharepoint 2007 in combinatie met Microsoft SQL Server 2005.

Uitgangssituatie

Data die nodig is voor het verkrijgen van het gewenste resultaat ligt opgeslagen in databases of bestanden of men moet in staat zijn deze data digitaal bij te houden. Verder moet de student toegang hebben tot de bronbestanden.

Benodigde software

Kantoorsoftware
Microsoft Sharepoint Server 2007
Microsoft SQL Server 2005
Microsoft Business Scorecard Builder

Benodigde hardware

Laptop of PC
Server

Beschikbare rapporten

Businessplan ETTU

Concrete werkzaamheden

Doorlopen van de zes fasen van een business intelligence project.

Uit te voeren activiteiten

Fase 1: Justification

Analyseren van de omgeving en de business
Risico's opstellen

Fase 2: planning

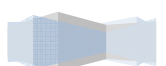
Hardware-eisen bepalen
Opzetten plan van aanpak
Uitgebreide planning opstellen

Fase 3: Business Analyses

Interviews proceseigenaren
Informatiebehoefte in kaart brengen
Succesfactoren bepalen
KPI's bepalen
Prioriteiten toekennen aan de KPI's
KPI's uitwerken
Systeemdokumentatie bekijken (indien aanwezig)
Bekijken of de informatiebehoefte kan worden voorzien mbv de huidige bronsystemen

Beschikbare tools bekijken

Bronanalyse Clockwise (urensysteem)
Bronanalyse Excel (budget)
Bronanalyse CRM (orderportefeuille)



Fase 4: Design

- Opstellen bronnenmodel
- Opstellen gegevensmodel
- Opstellen feitmodel
- Dimensie ontwerp
- Opstellen sterschema
- Opstellen ETL-schema
- Opstellen datamart

Fase 5: Construction

- ETL-development
- Application development

Fase 6: Deployment

- Implementation
- Testen
- Evaluation

Te hanteren methodieken

The complete project lifecycle for decision-support applications.

Te gebruiken technieken

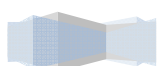
Kimball (datawarehouse), Sharepoint Portal Server 2007, interviewtechnieken

Resultaten voor de opdrachtgever (op te leveren produkten)

- Plan van aanpak (output fase 1/2)
- Business Requirements Report (output fase 3)
- Design report (output fase 4)
- Application (output fase 5)
- Evaluation (output fase 6)

Wat doet de opdracht uitstijgen boven een stageopdracht

Bij deze opdracht gaat het om het bedenken van een oplossing voor het voorzien in informatie. Het bepalen van informatiebehoefte, het analyseren van de omgeving en het bedenken en ontwikkelen van een werkende oplossing zijn vaardigheden die pas in hoofdfase twee worden aangeleerd. Ook is Business Intelligence een onderwerp dat pas na hoofdfase 1 wordt behandeld.



Bijlage B: Plan van aanpak

Inleiding

Dit plan van aanpak vormt het uitgangspunt voor het afstudeerproject van Ronald Kraijesteijn voor ETTU. In dit plan van aanpak zullen onderdelen worden opgenomen die van belang zijn voor het project, en vormt tevens een contract tussen opdrachtgever en opdrachtnemer.

Er zal worden begonnen met het beschrijven van de projectdefinitie waarin duidelijk zal worden waarom het project wordt uitgevoerd, welke fasen worden doorlopen en welke milestones er worden opgeleverd.

Vervolgens zal in de businesscase blijken waarom het noodzakelijk is om dit project uit te voeren. Hoofdstuk vier beschrijft de globale planning voor het project. Een gedetailleerde planning zal per week worden bijgehouden in een apart bestand. Tot slot zal de projectbeheersing worden beschreven.

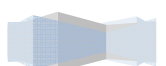
Projectdefinitie

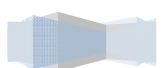
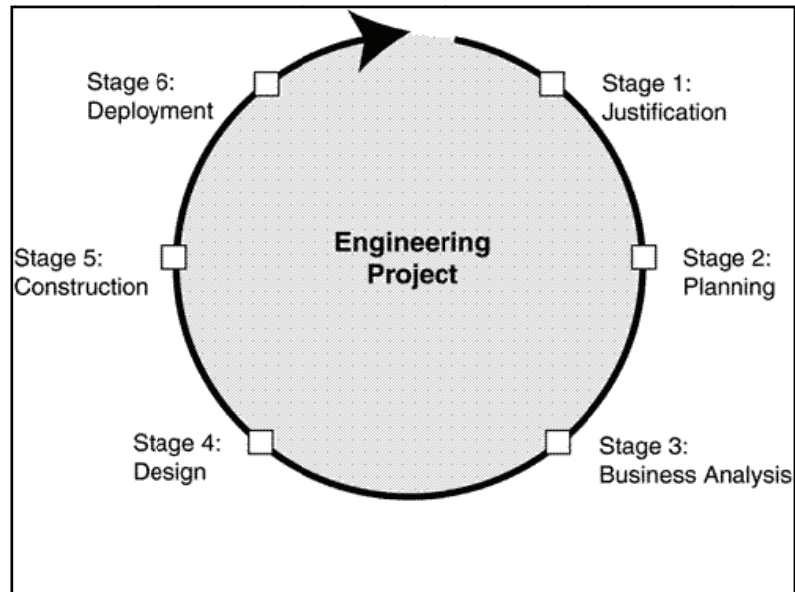
Doelstelling

Dit project heeft als doelstelling het opleveren van een Business Intelligence (BI) oplossing; een datawarehouse en een managementdashboard. Deze oplossing zal worden geïntegreerd binnen de sharepointomgeving van ETTU waardoor proceseigenaren in staat zijn hun proces beter te sturen.

Aanpak en fasering

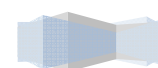
Voor een goede projectbeheersing is er gekozen om dit project in te delen in zes fasen. Iedere fase bevat een aantal activiteiten. Een grafische weergave van de fasen die tijdens dit project worden doorlopen is hieronder weergegeven. Vervolgens zullen ter verduidelijking de fasen worden beschreven.





Fasen en activiteiten

Fase	Activiteit	Toelichting/voorbeeld
1.0 Justification	Analyseren business	De BI-applicatie moet een zekere bijdragen leveren aan de bedrijfsdoelstellingen. Daarom is het noodzakelijk om zicht te krijgen op de business.
	Doelstelling van de Business Intelligence oplossing bepalen	Het is belangrijk om te weten welke doelen behaald willen worden door een BI-oplossing in te voeren.
	Projectrisico's vaststellen	Bekijken welke risico's er gelden voor dit project.
2.0 Planning	Technische zaken	Eisen hard- en software
	Niet technische infrastructuur bestuderen	Zijn er al datamodellen, wie zijn de stakeholders, is er documentatie over de systemen.
	Projectplanning	PID (projectplan) opstellen.
3.0 Business analyses	Functionele output	Welk soort informatie hebben medewerkers nodig, welke rapportages wil men graag zien, hoe moet dit eruit zien.
	Bronanalyse	Welke bronnen gaan we gebruiken, welke systemen, welke bestanden, welke databases, zijn er modellen van deze bronnen beschikbaar, wat is de kwaliteit van de data
4.0 Design	Datawarehouse modellering	Bronnenmodel, brongegevensmodel, feitenmodel, dimensietabellen, sterschema
	Front-end	Ontwerp schermen die de eindgebruiker ziet
	ETL	ETL-schema's, aggregatiemodellen
5.0 Construction	Datawarehouse	Database bouwen in sql server
	Front-end	Realisatie van de gewenste output.
6.0 Deployment	Testen	Systeem testen om de kwaliteit te controleren, plannen ETL proces.
	Support	Uitleg geven over de werking van het systeem.
	Invoering	Systeem overzetten van testomgeving in productieomgeving.



Milestones

In overleg met de opdrachtgever is er besloten (met het oog op de gekozen methode) de volgende milestones op te leveren:

Plan van aanpak

In dit document is uitgelegd hoe het project wordt aangepakt.

Business Requirements report

In dit document zal de informatiebehoefte vanuit de business worden beschreven. Onderdelen die in dit document worden opgenomen zijn:

- Succesfactoren, KPI's;
- Informatiebehoefte per proces;
- Bevindingen, problemen en aanbevelingen;
- Beschrijving bronsystemen en benodigde data;
- Beschrijving welke data gebruikt gaat worden;
- Beschrijving datakwaliteit;

Ontwerp

- Technische specificaties;
 - o Doeltabellen ontwerpen;
 - o Feitentabel en dimensietabel (sterschema);
 - o ETL-schema.
- Functionele specificaties
 - o Soort rapportage;
 - o Aantal rapportages;
 - o Lay-out van de rapportages;
 - o Schermontwerp;

Werkende oplossing

- Datawarehouse gebouwd in MS SQL SERVER 2005;
- Koppeling met Sharepoint portal;
- Gedefinieerde KPI's in Sharepoint, weergave in Sharepoint.
-

Afbakening

Kwaliteit

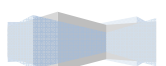
Om de kwaliteit te waarborgen zal het systeem worden getest in de laatste week van het project.

Budget

Omdat het om een afstudeerproject gaat is het budget niet van toepassing.

Mensen en middelen

Ronald Kraijesteijn zal het project zelfstandig uitvoeren. Om het project succesvol te kunnen voltooien zullen de medewerkers van ETTU bereid moeten zijn om mee te werken aan interviews en vragen over de bronsystemen en programmatuur.



Tijd

Dit project zal ongeveer 18 weken in beslag nemen, uitloop is niet mogelijk omdat het om een afstudeeropdracht gaat.

Scope

Maximaal drie KPI's weergeven per proces, uitgaan van de processen: bestuur, verkoop, planning en techniek.

Dit project zal zich beperken tot het gebruik van onderstaande informatiebronnen:

- Budgetsheet (Excel);
- Urenstelsysteem (Mysql);
- CRM (MS SQL).

Door de data te gebruiken uit deze drie bronsystemen kunnen er een aantal managementvragen worden beantwoord. In week vier zal er een definitief overleg worden gepleegd om te komen tot de juiste managementvragen. Het MOSCOW principe zal toegepast worden. Dit betekent dat we zullen kijken welke managementvragen in ieder geval beantwoord moeten worden en welke eventueel (wanneer er tijd over is) beantwoord zullen worden.

Methoden en technieken

De methode die zal worden gehanteerd: Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications (Larissa T. Moss, Shaku Atre). Niet alle activiteiten uit deze methode worden gebruikt. Deze methode zal worden gebruikt als een best practice.

Technieken die worden toegepast: Kimball (datawarehouse), SQL, Sharepoint.

Business Case

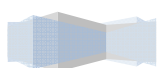
Aanleiding

ETTU is een jong bedrijf dat zich richt op de ontwikkeling van mobiele toepassingen. Verder ontwikkelt ETTU bedrijfsportals in Microsoft Sharepoint en levert consult op dit gebied. In de toekomst wil ETTU zich gaan richten op het detacheren van mensen.

Ettu heeft in het Businessplan het onderdeel "ADA factors" opgenomen (anders dan anderen). Hierin is bij het proces 'bestuur' de factor 'Real-time KPI presentatie' opgenomen. Door middel van deze factoren wil ETTU zich onderscheiden van de andere ICT-dienstverleners. Momenteel wordt er nog niet gewerkt met de Real-time KPI presentatie, welke proceseigenaren in staat stelt betere beslissingen te kunnen nemen.







ETTU werkt op uurbasis. Dit betekent dat er uren verkocht moeten worden richting de toekomst om zo de bezettingsgraad hoog te houden. Geen verkochte uren betekent geen werk voor het personeel. Een goede afstemming tussen orderportefeuille en beschikbare uren is gewenst.

ETTU verkoopt projecten. Voor een project is een budget ingepland die afgestemd is op het aantal benodigde uren. Het zou gewenst zijn wanneer ETTU real-time kon zien hoeveel uren er voor een project verbruikt zijn.



Risico's

Dit project brengt een aantal risico's met zich mee. Onderstaande tabel toont een overzicht met alle risico's die voor dit project van toepassing zijn.

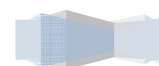
Tabel 1: Risicotabel			
Risico	Omschrijving	Risico	Maatregel
Technologie	Weinig ervaring met de technologie		Zelfstudie, goede inzet
Planning	Te krap plannen, project niet op tijd klaar		Planning in deelfasen, werk beter inschatten, beperken tot de primaire processen
Integratie	Integreren binnen SharePoint geeft problemen		Hulp inschakelen van expert op dit gebied
Support	Weinig kennis binnen de organisatie op het gebied van business Intelligence		Goede zelfstandigheid van de uitvoerder, support op school vragen
Communicatie over en weer	Te weinig interactie kan ervoor zorgen dat de wens van de opdrachtgever niet goed is afgestemd met het eindresultaat		Wekelijks vergadermoment inbouwen in de planning
Datakwaliteit	Kwaliteit data slecht of niet integer, daardoor kloppen de rapportages niet		Bewust maken van het belang van goede datakwaliteit, controleren op kwaliteit

 Gemiddeld risico

 Hoog risico

Succesfactoren

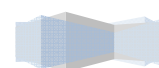
Het op tijd opleveren van het project in verband met de beperkte tijd die ervoor staat. Het niet op tijd opleveren kan ertoe leiden dat het project mislukt. Verder is de kwaliteit van de brondata een cruciale factor voor het slagen van het project.

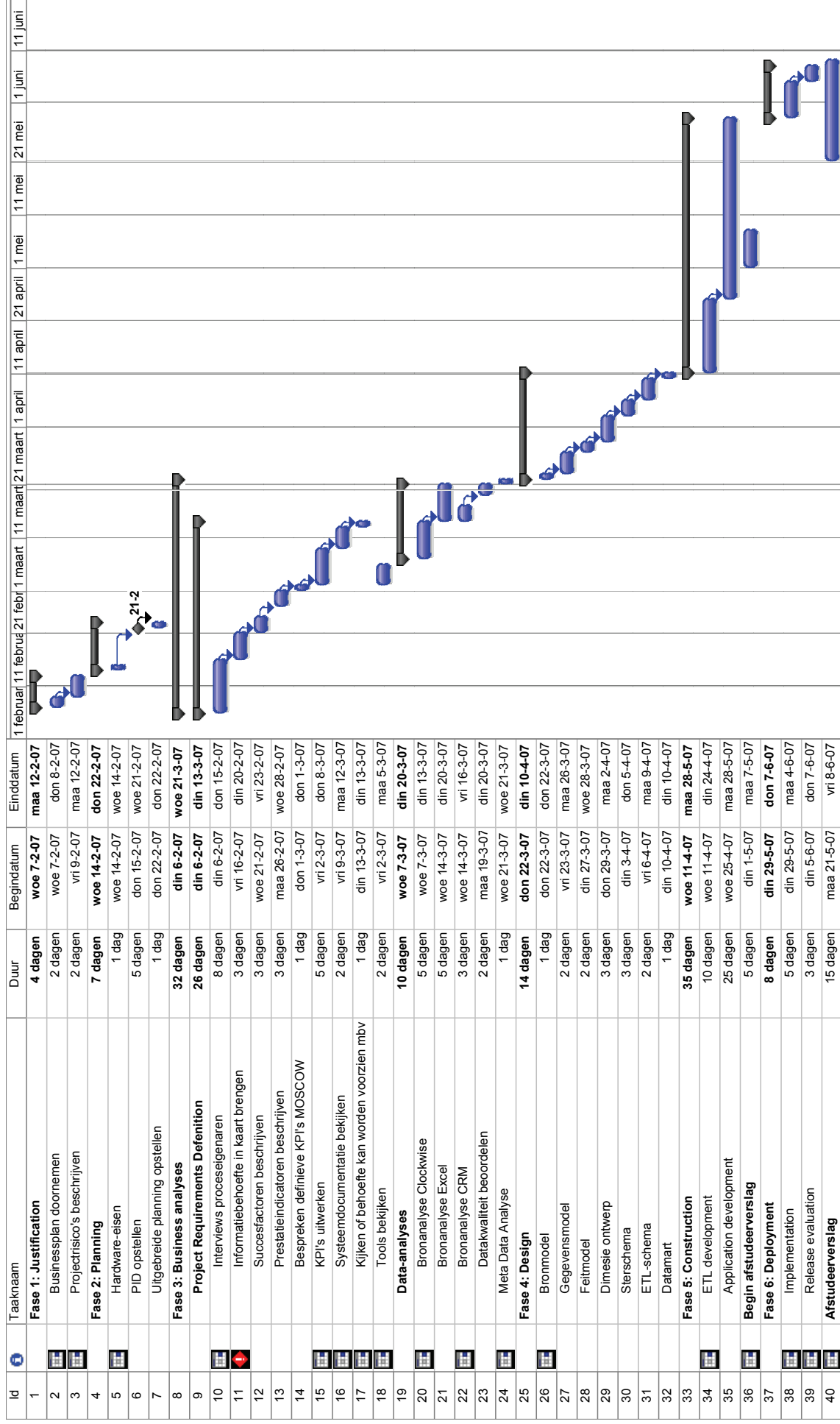


Planning

Fase	Geplande startdatum	Geplande einddatum
1.0 Justification	5-2-2007	12-2-2007
2.0 Planning	14-2-2007	22-2-2007
3.0 Business analyses	6-2-2007	21-3-2007
4.0 Design	22-3-2007	10-4-2007
5.0 Construction	11-4-2007	28-5-2007
6.0 Deployment	29-5-2007	8-6-2007

Op de volgende pagina is een detailplanning te vinden van de fasen en de activiteiten.





Project beheersing

Voortgangsrapportages

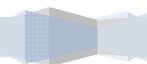
Ieder week zullen er een of twee gesprekken worden gehouden met de opdrachtgever om de voortgang te bewaken. Verder zullen alle projectdocumenten op het portal van ETTU beschikbaar worden gesteld zodat de stakeholders kunnen zien hoe het project vordert.

Voorwaarden aan derden

De enige partij die is betrokken naast ETTU is de Haagse Hogeschool. De Haagse Hogeschool zal Ronald Kraijesteijn op de hoogte brengen van terugkomdagen, nog te volgen workshops en deadlines. De Haagse Hogeschool zal een rapport opgeleverd krijgen waarin is te lezen welke activiteiten tijdens dit project zijn uitgevoerd, aangevuld met bijlagen van de deelproducten.

Communicatie en vastlegging

Communicatie zal plaatsvinden tussen de opdrachtgever en opdrachtnemer, zowel mondeling in de vorm van vergaderingen als schriftelijk via e-mail. Overige communicatie met medewerkers zal zoveel mogelijk mondeling worden verricht. Er zullen van iedere vergadering notulen worden bijgehouden en op het portal worden geplaatst.



Bijlage C: Business Requirements Report

Inleiding

Business intelligence draait om het nemen van beslissing aan de hand van samengevatte rapportages en key performance indicatoren in de vorm van metertjes of lampjes. Alvorens te komen tot de juiste key performance indicatoren is het belangrijk te weten welke zaken belangrijk zijn om te meten. Het business requirements report is een document dat een overzicht geeft van de informatiebehoefte die op dit moment speelt. Dit document geeft antwoord op onderstaande vragen:

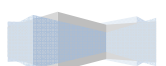
- Welke informatie is er nodig?
- Hoe moeten de rapportages er globaal uitzien?
- Hoe schoon is de data en waar moet op gelet worden?

Het hoofddoel van dit document is om in kaart te brengen wat de informatiebehoefte van ETTU is. ETTU verkoopt uren, en om de juiste afstemming te vinden tussen bezetting en urenverkoop is het van belang in de toekomst vooruit te kunnen kijken. Ook acteren mensen op de wijze waarop zij gemeten worden wat een positief resultaat teweeg kan brengen.

Nadat de informatiebehoefte is beschreven zal er als eerste worden gekeken wat de visie is van ETTU. Daaruit kunnen vervolgens de succesfactoren worden afgeleid. Vervolgens zal er worden gekeken welke succesfactoren het meest belangrijk zijn. Voor deze succesfactoren zullen de KPI's worden opgesteld nadat er met de betreffende proceseigenaar is overlegd.

Om een beter beeld te kunnen vormen van de systemen en data is er gekeken naar de bronsystemen die nodig zijn voor de analyse. De volgende activiteiten zijn uitgevoerd:

- Welke gegevens zijn (tenminste) nodig voor de analyse;
- Wat is de kwaliteit;
- Waar moet op gelet worden.



Informatiebehoefte per proces

Inleiding

In dit hoofdstuk wordt de informatiebehoefte per proces beschreven. Deze behoefte is beschreven nadat er met iedere proceseigenaar een interview is gehouden in de weken een en twee van dit project. Daarin hebben zij aangegeven welke informatie zij graag overzichtelijk op het scherm willen hebben in de vorm van een managementdashboard. Onderstaand overzicht toont de belangrijkste managementvragen per proces. Er zijn diverse managementvragen die meerdere keren naar voren kwamen in de gesprekken zoals het gemiddeld verkooptarief of de bezettingsgraad. In het volgende hoofdstuk is een tabel te vinden met alle KPI's. Daar is ook in beschreven welke processen behoefte hebben aan deze KPI's.

De hierop volgende paragrafen laten zien welke vragen er op dit moment spelen bij de diverse proceseigenaren. In hoofdstuk drie zullen deze vragen vertaald worden in meetbare eenheden, de KPI's (key performance indicators).

Informatiebehoefte proces verkoop

1. Ik wil graag real-time kunnen zien hoeveel geld er aan orders in de orderportefeuille aanwezig is. Dit moet dan naast de forecast gezet worden om te zien of de forecast gehaald gaat worden.
2. Ik wil graag kunnen zien hoeveel nieuwe klanten er per maand zijn bijgekomen.
3. Ik wil een aantal algemene verkoopgegevens zien zoals het gemiddelde uurtarief en het bezettingspercentage per medewerker, omzet per klant.

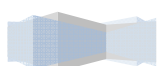
Informatiebehoefte proces bestuur

1. Ik wil graag een overzicht van diverse zaken hebben die met bestuur te maken hebben zoals bezettingsgraad (real-time), uurtarieven(real-time), etc. Deze wil ik afgezet zien tegen het budget zodat ik weet hoe we ervoor staan of hoe de maand is gelopen.

Informatiebehoefte proces techniek

1. Ik wil graag real-time kunnen zien hoeveel geld er aan orders in de orderportefeuille aanwezig is. Dit moet dan naast de forecast gezet worden om te zien of de forecast gehaald gaat worden.
2. Evenwicht tussen de voorraad uren en de benodigde uren die in de planning zijn vastgelegd.

Ik wil graag real-time kunnen zien hoeveel budget er voor een project is verbruikt. Eventueel is dit ook handig voor de mensen op de werkvloer.



Prestatie-indicatoren en succesfactoren ETTU

Inleiding

Business Intelligence draait om het nemen van de juiste beslissingen om strategische doelen te kunnen bereiken. Het meten van succesfactoren kan met behulp van prestatie-indicatoren. De eerste stap bij het vertalen van strategische doelen in prestatie-indicatoren is het vertalen van strategische doelen (bijv. continuïteit, winst, groei) naar kritische succesfactoren.

Strategische doelen

1. Groei
2. Projecten verkopen

Kritische succesfactoren ETTU

Kritische succesfactoren geven aan op welke gebieden ETTU moet excelleren om de strategische doelen te realiseren. Deze factoren zijn afgeleid uit het missiestatement welke staat omschreven in het businessplan. Verder zijn er een aantal KSF's benoemd in het businessplan.

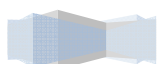
Onderstaand overzicht toont alle succesfactoren van de firma ETTU.

Algemene KSF

1. Betrouwbaar zijn
2. Winstgevend zijn
3. Klantgericht zijn
4. Uitstekende reputatie hebben
5. Leveren van consultancy en development
6. Hoog opgeleid personeel hebben
7. Personeel zich thuis laten voelen
8. Open, transparant en innovatief zijn
9. Garant staan voor de kwaliteit van opgeleverde oplossingen
10. Investeren van de klant waarborgen

KSF Verkopen

11. Goede opvolging van afspraken
12. Snelle response van techniek
13. Professionalisering en standaardisering van PVA en offertes detachering
14. Meer inzicht in prospectieve situatie
15. Inzet van jr account manager
16. Goede coaching van jr. account manager door SU



KSF Techniek

1. Goed functioneren Total Quality Management
2. Projectvolgving perfect

KSF Personeelszaken en werving

1. Adverteren
2. Inhoudelijk mentorschap

KSF Bestuur

1. Managementdevelopment
2. Ondersteuning segment en Ceasar organisatie

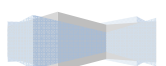
In de volgende paragraaf zal worden besproken hoe succesfactoren kunnen worden gemeten.

Keuze voor uitwerking Key Succes Factoren

Succesfactoren kunnen worden gemeten door een of meerdere prestatie-indicatoren. Key performance indicators zijn 'metertjes' waarmee een belangrijk aspect van de resultaten wordt gemeten. Deze indicator hoort een goede indicatie te geven voor succes of falen op het gebied van een succesfactor. Een key performance indicator is geen doel of middel op zich, maar een objectieve manier om na een bepaalde periode terug te kijken naar de prestaties ten behoeve van een doel of succesfactor. Een key performance indicator zorgt er dus voor dat belangrijke zaken tijdig onderkend worden waardoor er tijdig actie kan worden ondernomen. Bij het bepalen van goede indicatoren zal er rekening worden gehouden met de focus van de strategische doelen.

Voor dit project ligt de focus op een beperkt aantal succesfactoren. Er wordt gefocust op KSF's die betrekking hebben op de primaire processen van ETTU. Dit zijn ook de KSF's die het management het belangrijkste vindt om te meten. Er is gekozen om de volgende succesfactoren uit te werken:

- Projectvolgving;
- Winstgevend zijn;
- Klantgericht zijn;
- Meer inzicht in prospectieve situatie;



KPI's per proces

Onderstaande tabel toont alle potentiële KPI's die voor dit project in aanmerking komen. Omdat niet alle KPI's even belangrijk zijn zal er een afweging worden gemaakt welke KPI's de voorkeur genieten en welke minder belangrijk zijn.

KPI	Proces	Weergave
Opslagkosten (maand,jaar,BU,totaal)*	Bestuur	Piet
Bruto toegevoegde waarde (maand, BU, totaal)*	Bestuur	Piet
Bedrijfsresultaat (maand, BU, totaal)*	Bestuur	Piet
Directe kosten per maand (BU, totaal)	Bestuur	Piet
Winst per project	Techniek	Patrick, Eric-Jan, Piet
Percentage gebruikt budget per project	Techniek	Patrick / Piet
Ziektepercentage (medewerker,BU,jaar,maand)	Techniek	Piet, Patrick
Verlofpercentage (medewerker, BU)	Techniek	Piet
Gemiddelde omzet per klant	Verkoop	Piet / Vincent
Gemiddeld verkooptarief (maand, jaar, BU, medewerker)*	Verkoop	Piet, Patrick, Vincent
Bezettingsgraad (BU, medewerker)	Verkoop	Piet, Patrick
Orderportefeuille per maand*	Verkoop	Piet, Eric-Jan,Vincent
Werkelijke omzet*	Verkoop	Patrick
Aantal nieuwe klanten per maand	Verkoop	Piet, Vincent

*Notities

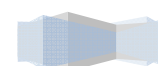
Opslagkosten per maand: als percentage vd bruto toegevoegde waarde
 Bruto toegevoegde waarde: afgezet tegen het budget
 Bedrijfsresultaat: afgezet tegen het budget
 Directe kosten per maand: als percentage vd bruto toegevoegde waarde
 Gemiddeld verkooptarief: afgezet tegen het budget
 Orderportefeuille: als percentage vd omzet uit de forecast
 Werkelijke omzet: als percentage van het budget

Maandag 26 februari 2007 zijn de KPI's gepresenteerd aan de proceseigenaren. Vervolgens is er in overleg met Vincent Heijckmann en Piet Vink besloten om een prioriteitenlijst te maken omdat er beperkte tijd beschikbaar is voor het project.

Het weergeven van onderstaande KPI's geniet de voorkeur:

1. Opslagkosten (indirecte kosten)
2. Directe kosten
3. Bruto toegevoegde waarde
4. Orderportefeuille
5. Gebruikt budget per project

Er zal dan ook getracht worden om tenminste deze vijf KPI's weer te geven in het uiteindelijke managementdashboard. Onderstaande tabel toont een overzicht van de KPI's die in dit project uitgewerkt gaan worden. In de eerste kolom is te zien welk proces verantwoordelijk is voor de KPI, daarna is af te lezen welke succesfactor ermee gemoeid is en tot slot de KPI zelf in de laatste kolom.



Proces	Succesfactor	Prestatie-indicator
Bestuur	Winst maken	Percentage maximale directe kosten
		Percentage maximale indirecte kosten
		Bruto toegevoegde waarde
Techniek	Projectvolging	Verbruikt budget per project
Verkoop	Prospectieve situatie	Orderportefeuille per maand
	Klantgericht zijn	Verbrand budget per project

Uitwerken van de KPI's

Inleiding

Zoals eerder is beschreven is er besloten een beperkt aantal succesfactoren met bijhorende KPI's mee te nemen in dit project. Het doel van dit hoofdstuk is om de KPI's verder uit te werken. De KPI's zijn uitgewerkt naar proces. Dit is gedaan om het geheel overzichtelijker te maken.

Per KPI is er gekeken naar onderstaande zaken:

- **Omschrijving:** omschrijving van de KPI
- **Doel:** waarom is het nuttig om deze KPI weer te geven
- **Berekening:** hoe kan de KPI ontstaan
- **Bron:** waar komen de gegevens vandaan
- **Streefwaarde:** welke waarde wil men bereiken, ook wel 'goal' genaamd
- **Frequentie:** hoe vaak moet de KPI ververs worden
- **Uitvoer:** hoe ziet de (mogelijke) uitvoer eruit

KPI's proces 'bestuur'

KPI Opslagkosten

Omschrijving: opslagkosten (indirecte kosten)

Doel: het doel van deze KPI is om overzichtelijk te maken wat de percentages indirecte kosten zijn.

Berekening: opslagkosten / BTW * 100%








Bron: excelsheet

Streefwaarde: <=40%

Frequentie: einde van de maand 1x (na het invoeren van de werkelijke gegevens)

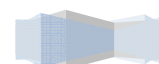
Data nodig

- Opslagkosten per businessunit, per maand
- BTW per businessunit, per maand

	Januari	Februari	Maart
+Totaal	40% 
-Applications	35% 	35% 	36% 
-Configurations	45% 	39% 	38% 

Hoe werkt het

Begin van het jaar wordt het budget vastgesteld. Dit budget is opgesplitst in businessunits;



Applications , Configurations en Detachering. In het budget is vastgelegd wat er voor winst behaald moet worden, welke kosten er zijn, etc. Aan het einde van iedere maand worden de werkelijke cijfers in het bestand ingevoerd. Het is de bedoeling dat in een oogopslag duidelijk wordt of de indirecte kosten niet boven de 40% stijgen. Dit kan door middel van het weergeven van een KPI die de opslagkosten uitrekent als percentage van de bruto toegevoegde waarde.

Issue: data staat in Excel en zal moeten worden getransferd naar SQL server.

KPI directe kosten

Omschrijving: Directe kosten per maand laten zien als percentage vd BTW.








Doel: het doel van deze KPI is om overzichtelijk te maken wat de percentages directe kosten zijn.

Berekening: directe kosten/BTW*100%

Bron: excelsheet

Streefwaarde: **nader te bepalen**

Frequentie: 1x per maand

	Januari	Februari	Maart
+Totaal	40% 
-Applications	35% 	35% 	36% 
-Configurations	45% 	39% 	38% 

KPI bruto toegevoegde waarde

Omschrijving: BTW laten zien als percentage van de gebudgetteerde BTW.

Doel: btw afzetten tegen het budget, kijken of er boven budget is gepresteerd

Berekening: BTW behaald / BTW budget * 100% = Percentage




Bron: excelsheet

Streefwaarde: 100% of meer

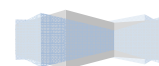
Frequentie: einde van de maand 1x

Data nodig

- BTW per businessunit, per maand
- Gebudgetteerde BTW per businessunit, per maand

	Januari			...
	Actual	Target	%	
+BTW Totaal	1000	2000	50% 	
-BTW Applications	500	1000	50% 	
-BTW Configurations	500	1000	50% 	

Notes: voor de aankomende maanden zal het budget en de forecast BTW met elkaar overeenkomen. Dit zegt dus niet veel. Als de maanden verstreken zijn en de werkelijke cijfers zijn ingevoerd kloppen de waarden.



KPI's proces 'techniek'

KPI verbruikt budget per project

Omschrijving: percentage verbruikt budget laten zien voor een project (in geld).

Doel: deze KPI laat zien hoeveel budget er voor een project verbruikt is. Als blijkt dat tegen het einde van het project het budget bijna op is kan er bijgestuurd worden, bijvoorbeeld door extra capaciteit in te schakelen. Eventueel kan deze KPI ook geïntegreerd worden in het portal zodat de klanten kunnen zien hoeveel uren er aan hun project is verspijkerd.

Berekening: verbruikt: uren geschreven voor een project per persoon * uurtarief binnen dit project.
Maximaal budget: is vastgelegd in het systeem bij het desbetreffende project.

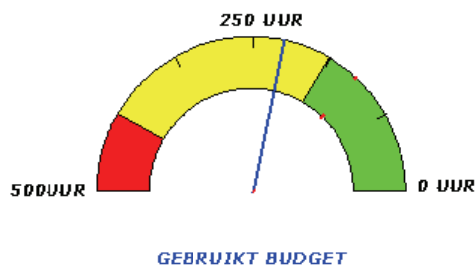
Bron: clockwise (urensysteem)

Streefwaarde: vooraf bepaalde fixed price die moet worden vastgelegd per project

Frequentie: wekelijks

Data nodig

- geschreven uren per project (som uren uit u_uren, per project);
- maximaal te gebruiken budget per project;
- Uurloon van de medewerker die uren schrijft voor het project.



n.b. uren worden bedragen

Hoe werkt het

Er wordt in clockwise ingevoerd hoeveel geld er aan een project mag worden besteed. Vervolgens worden er een aantal medewerkers aan het project gekoppeld. Deze kunnen dan ook uren voor het betreffende project schrijven.

In het schermplje "maandoverzicht" binnen clockwise is er per medewerker te zien hoeveel uur zij deze maand hebben geschreven. Het aantal geschreven uren per project wordt berekend door het totale aantal uren dat geschreven is voor het project, ongeacht of de uren goedgekeurd zijn of niet.

Issue: het zou mooier zijn om het aantal gebruikte uren in een metertje te zetten omdat het budget een fixed price is die niet altijd op het aantal uren is gebaseerd.

KPI's proces 'verkoop'

KPI prospectieve situatie

Omschrijving: Orderportefeuille laten zien, afgezet tegen de forecast

Doel: duidelijk maken of er genoeg orders zijn voor een maand

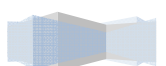
Berekening: som orderportefeuille per maand nemen en dit delen afzetten tegen de forecast omzet

Bron: CRM, excelsheet





Streefwaarde: 1: 100%, 2: 80%, 3: 40%, 4: 20%

Dimensies: maand

Frequentie: begin van de maand, 1x



Uitvoer:

Maand	Orders	Forecast	%	Status
1	25.000	25.000	100%	
2	80.000	10.000	80%	
3	10.000	30.000	33%	
4	2.000	20.000	10%	

Datakwaliteit

Inleiding

Om te komen tot de juiste beslissingen is het van essentieel belang dat de kwaliteit van de data integer is. De kwaliteit is integer wanneer de data geschikt is voor het doel waarvoor het gebruikt wordt. Wanneer dat niet het geval is zal dit ertoe leiden dat de KPI's die uiteindelijk worden weergegeven niet kloppen. Wanneer de KPI's niet kloppen betekent dit dat het project mislukt. In dit hoofdstuk zal er worden beschreven welke maatregelen er zijn genomen of zullen worden genomen om de kwaliteit van de data te vergroten. Risico's die worden gelopen bij het overslaan van deze stap:

- Verkeerde beslissingen nemen;
- Te laat beslissingen nemen;
- Frustratie; fouten lastig te achterhalen;
- Strategiebepaling kost meer tijd of wordt niet gerealiseerd.

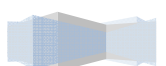
Kortom, het bekijken en verbeteren van de datakwaliteit is belangrijk. Voor het beoordelen van datakwaliteit wordt er gelet op onderstaande punten:

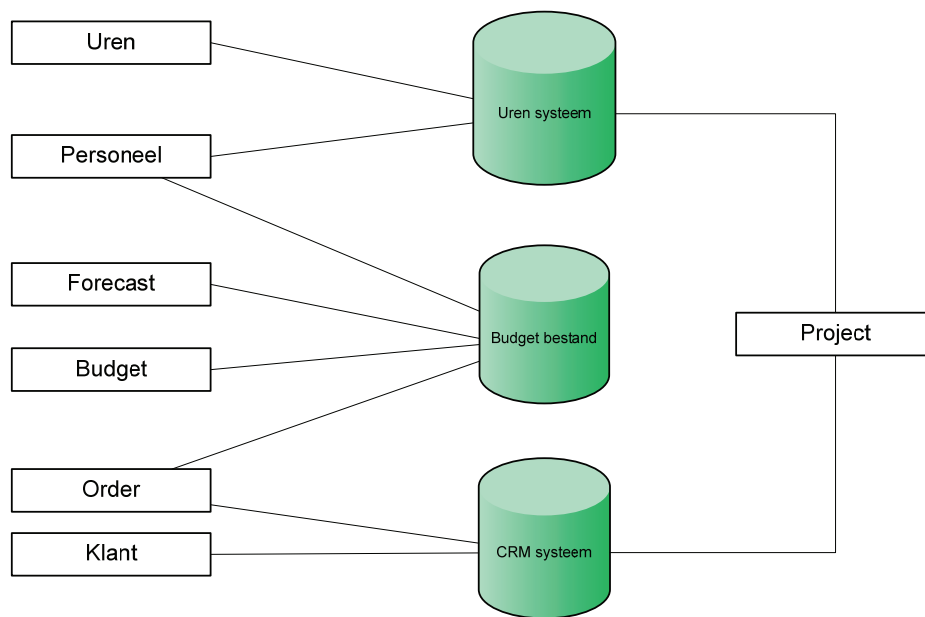
- Integriteit;
- Tijdigheid;
- Volledigheid;
- Consistentheid;

Per bronsysteem is er gekeken wat de kwaliteit is, gelet op bovenstaande punten.

Bronsystemen

Onderstaand bronnenmodel toont de drie bronnen die we voor dit project gaan gebruiken. De diverse bronnen gebruiken enkele overlappende gegevens zoals personeelsinformatie en projectinformatie.





De drie bronnen zijn geen van allen van hetzelfde type. Het urensysteem maakt gebruik van een MySQL database, het budgetbestand is een spreadsheet en het CRM systeem is een MS SQL database. Dat betekent dat de gegevens door middel van een ETL proces omgezet moeten worden naar een formaat database.

Datakwaliteit Excel

Integriteit

De proceseigenaar van het proces 'bestuur' zal er op moeten letten dat de data die hij invult integer is. Hij zal dus bij het invullen van het budget/forecast alert moeten zijn op het feit dat er fouten kunnen ontstaan in de rapportages wanneer de data die hij invult niet klopt.

Tijdigheid

Er kan worden uitgegaan dat de excelsheet altijd beschikbaar is. Deze zal worden opgeslagen op het netwerk op dezelfde server als waar het datawarehouse draait. Men zal er voor moeten zorgen dat deze file op een en dezelfde locatie wordt opgeslagen.

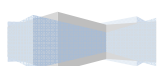
Volledigheid

Het is noodzakelijke dat alle velden die nodig zijn om de gewenste rapportage uit te draaien ingevuld zijn. De volgende velden dienen ten minste ingevuld te zijn:

- Forecast BTW alle maanden, beide businessunits
- Forecast opslagkosten, alle maanden, beide businessunits
- Forecast directe kosten, alle maanden, beide businessunits

Consistentheid

Niet van toepassing



Datakwaliteit Clockwise

Integriteit

Personeel moet goed controleren of de ingevulde uren kloppen omdat er geen check op het programma zit. Iemand kan bijvoorbeeld 100 uur op een dag schrijven.

Tijdigheid

De database van clockwise draait op een Mysql server. Deze server dient beschikbaar te zijn voor toegang op het moment dat het datawarehouse ververs wordt. We kunnen er vanuit gaan dat de database 24x7 up and running is. Personeel moet tijdig de uren invullen. Leidinggevenden moeten tijdig de uren accepteren.

Volledigheid

De urenverantwoording moet iedere vrijdag voordat het personeel naar huis gaat ingevuld zijn. Verder moeten alle dagen van de verstreken week ingevuld zijn. Verder moeten de volgende gegevens ingevoerd zijn:

- Uurtarief per medewerker specifiek per project
- Fixed prijs per project
- Volledige urenverantwoording per medewerker, per week

Consistentheid

Er moet gecontroleerd worden of ieder project slechts een maal in het systeem staat. Een aantal projecten zijn dubbel opgeslagen onder dezelfde naam.

Datakwaliteit CRM

Integriteit

Bij het vullen en updaten van het CRM systeem moet de volgende data kloppen: maand waarin de order valt, orderbedrag en het slagingspercentage.

Tijdigheid

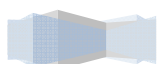
De database van CRM draait op een SQL server. Deze server dient beschikbaar te zijn voor toegang op het moment dat het datawarehouse ververs wordt. We kunnen er vanuit gaan dat de database 24x7 up and running is.

Volledigheid

De data moet in ieder geval volledig zijn aan voor het einde van de maand.

Consistentheid

NVT



Bijlage D: Business Design Report

Inleiding

Het ETTU dashboard is niet zomaar ontstaan. Om de informatie binnen dit dashboard weer te geven zijn er een aantal stappen doorlopen. Een van die stappen is het ontwerpen van schema's waaronder die van het datawarehouse. Dit document bevat informatie over de opbouw en de werking van het datawarehouse. Het moet een beter beeld geven over welke stappen zijn doorlopen om het uiteindelijke managementdashboard van ETTU te vullen. Mochten er aanpassing nodig zijn binnen het dashboard kan dit document geraadpleegd worden. Er staat namelijk in beschreven welke bestanden, tabellen en records er gebruikt worden.

Onderstaande stappen zullen worden doorlopen om tot een goed ontwerp te komen.

1. **Bronnenmodel opstellen**

Het eerste model betreft een overzicht van de beschikbare bronnen. Elk tonnetje bevat een database of spreadsheet. In de rechthoeken staan de gegevensgebieden. Met behulp van de gegevensgebieden wordt aangegeven welke overlap/aanknopingspunten er tussen de bronnen bestaan en wat de kern van de gegevens per bron is. De gegevensgebieden komen niet met een tabel overeen maar met een hele groep tabellen.

2. **Brongegevensmodel**

Om meer inzicht te krijgen op de inhoud en samenhang van de bronnen gebruiken we een relationeel gegevensmodel van de gezamenlijke bronnen.

3. **Feiten- dimensiemodel**

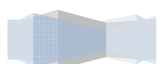
Een feitentabel definieert de verschillende toestanden of gebeurtenissen. Bijvoorbeeld het aantal gebruikte uren voor een project. Dimensies worden gebruikt om feiten te categoriseren zodat ze overzichtelijk worden gepresenteerd. Elke dimensie beschrijft een invalshoek van de feiten.

4. **Sterschema**

In dit ontwerp (sterschema) worden de associaties tussen tabellen duidelijk: van een feitentabel naar een dimensietabel.

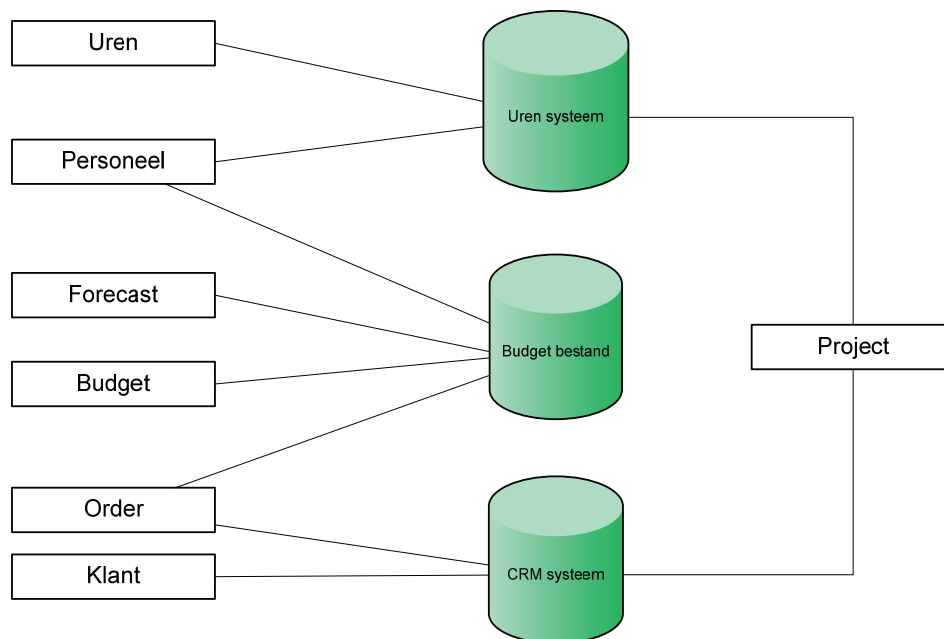
5. **ETL-schema's**

Niet alle gegevens komen uit dezelfde database. Tussen de diverse gegevensbronnen zitten gegevens die elkaar aanvullen. Alle gegevens die we willen betrekken bij de analyse zullen we moeten samenbrengen in een datawarehouse. Het ETL-schema beschrijft de relatie tussen brontabellen en doeltabel in het datawarehouse, met alle noodzakelijke bewerkingen.

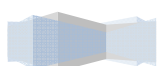


Opstellen bronnenmodel

Onderstaand bronnenmodel toont de drie bronnen die we voor dit project gebruiken. De diverse bronnen gebruiken enkele overlappende gegevens zoals personeelsinformatie en projectinformatie. Of deze overlappende informatie ook gebruikt zal worden is nog niet duidelijk.



De drie bronnen zijn geen van allen van hetzelfde type. Dit is belangrijk om te weten omdat de data straks moet worden overgezet in een datawarehouse die gebruikt maakt van een MS SQL SERVER database. Het urensysteem (clockwise) maakt gebruik van een MYSQL database, het budgetbestand is een spreadsheet en het CRM systeem is een MS SQL database. .



Opstellen brongegevensmodellen

De data die we nodig hebben om de rapportages te kunnen vormen is opgeslagen in drie verschillende bronsystemen, namelijk Mysql, Excel en Microsoft SQL Server 2005. Dit hoofdstuk beschrijft welke velden we gaan gebruiken in het datawarehouse.

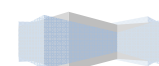
De data die in Excel is opgeslagen is verdeeld over verschillende tabbladen. Onderstaande tabellen laten zien welke data er per tabblad benodigd is. Uiteindelijk zullen we een nieuw tabblad aanmaken waarin alle benodigde data is te vinden. Dit is nodig voor een goede transfer naar het datawarehouse.

Brongegevensmodel budgetbestand

Het budgetbestand hebben we nodig om KPI's te kunnen weergeven voor de BTW, directe kosten, indirecte kosten en de orderportefeuille. Dit bestand is gemaakt in Excel en bestaat uit verschillende tabbladen van waaruit we gegevens nodig hebben.

De eerste drie tabbladen van waaruit we gegevens nodig hebben zijn die van de forecast per businessunit. Onderstaande tabellen tonen welke velden we gebruiken.

Bestand	Attribuut	DataType	Omschrijving
Forecast applications	Maand	String	Maand
	Subtotaal opslagkosten	Integer	Som kosten
	BTW	Integer	Bruto Toegevoegde waarde
	Totale Dir Pers kstn	Integer	Directe kosten
Bestand	Attribuut	DataType	Omschrijving
Forecast configurations	Maand	String	Maand
	Subtotaal opslagkosten	Integer	Som kosten
	BTW	Integer	Bruto Toegevoegde waarde
	Totale Dir Pers kstn	Integer	Directe kosten
Bestand	Attribuut	DataType	Omschrijving
Forecast detachering	Maand	String	Maand
	Subtotaal opslagkosten	Integer	Som kosten
	BTW	Integer	Bruto Toegevoegde waarde
	Totale Dir Pers kstn	Integer	Directe kosten

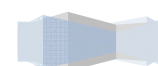


De overige tabbladen die we gebruiken zijn die van het budget. Per businessunit is er een tabblad waarin het budget is vastgelegd voor het huidige jaar.

Bestand	Attribuut	DataType	Omschrijving
Budget applications	Maand	String	Maand
	Bruto toegevoegde waarde	Integer	BTW
	Kostprijs directe medewerkers	Integer	Directe kosten
	Totaal opslagkosten	Integer	Indirecte kosten

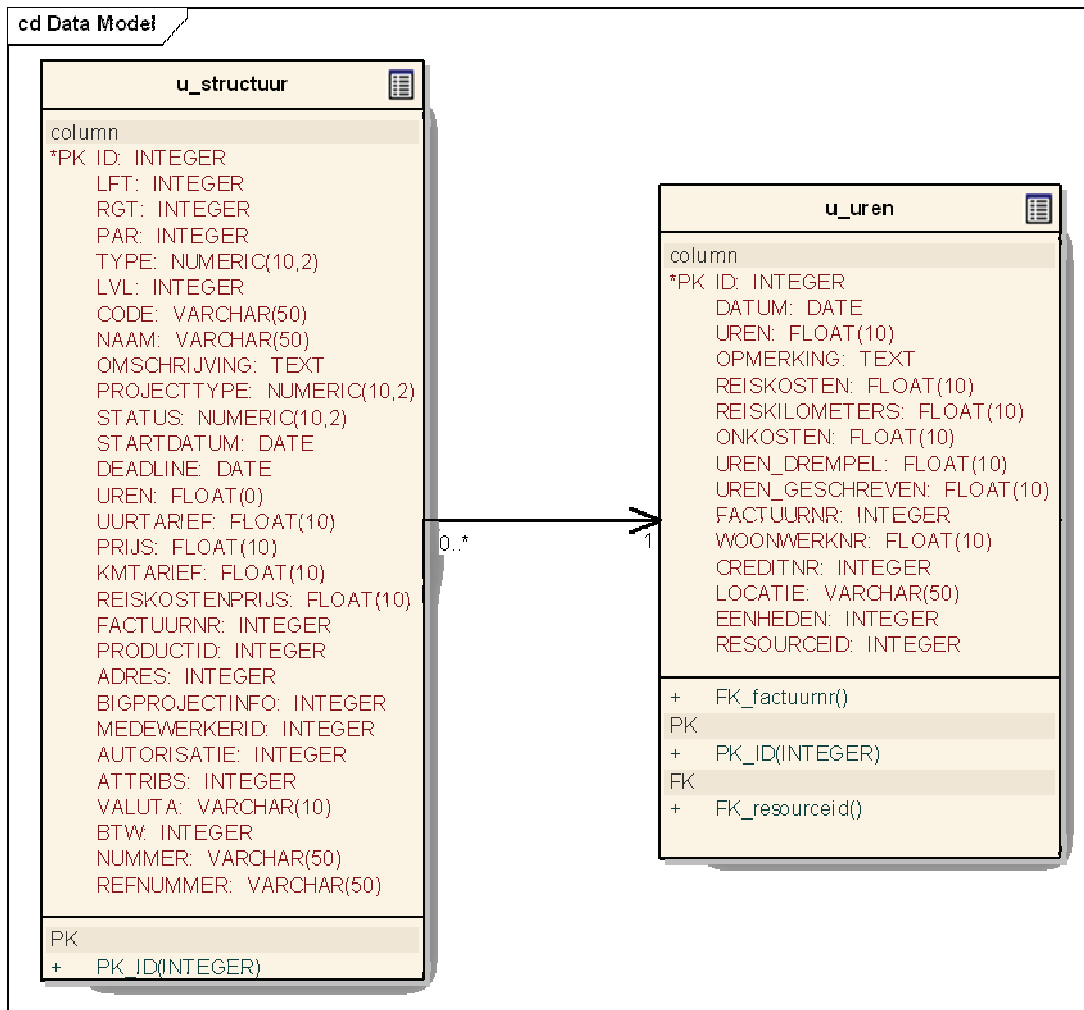
Bestand	Attribuut	DataType	Omschrijving
Budget configurations	Maand	String	Maand
	Bruto toegevoegde waarde	Integer	De som
	Kostprijs directe medewerkers	Integer	Directe kosten
	Totaal opslagkosten	Integer	Indirecte kosten

Bestand	Attribuut	DataType	Omschrijving
Budget deta	Maand	String	Maand
	Bruto toegevoegde waarde	Integer	De som
	Kostprijs directe medewerkers	Integer	Directe kosten
	Totaal opslagkosten	Integer	Indirecte kosten



Brongegevensmodel urensysteem

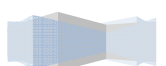
Het urensysteem, ook wel bekend als “clockwise” is een systeem dat gebruikt maakt van een MySQL database en welke bestaat uit ongeveer 15 tabellen. Om het gebruikte budget per project te kunnen berekenen hebben we onderstaande twee tabellen nodig.



De enige twee tabellen die we nodig hebben zijn de tabellen **u_structuur** en **u_uren**. In de **u_structuur** tabel staan de projecten vastgelegd en de medewerkers die aan een project zijn gekoppeld. Het veld “type” bevat respectievelijk de waarde K, P of R (klant/project/resource). Voor iedere gekoppelde medewerker is er dus een **u_structuur** record waarbij PAR bepaald aan welk project de medewerker gekoppeld is.

In de tabel **u_uren** staan de geschreven uren. Deze uren zijn op hun beurt gekoppeld aan een medewerker en project door middel van de sleutel **RESOURCEID**. Dit is het veld ID uit **u_structuur** met record TYPE R.

Brongegevensmodel CRM-systeem

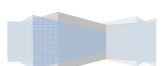


De database van CRM hebben we nodig om de orderportefeuille per maand te tonen. De database die voor CRM gebruikt wordt is erg groot. Voor onze rapportage hebben we slechts een tabel nodig, namelijk: OpportunityBase. Onderstaand gegevensmodel toont hoe de tabel is opgebouwd.

OpportunityBase	
column	
*PK OpportunityId: uniqueidentifier	
* DeletionStateCode: int	
PriceLevelId: uniqueidentifier	
OpportunityRatingCode: int	
PriorityCode: int	
ContactId: uniqueidentifier	
AccountId: uniqueidentifier	
Name: nvarchar(300)	
StepId: uniqueidentifier	
Description: ntext	
EstimatedValue: money	
Stepname: nvarchar(200)	
SalesStageCode: int	
ParticipatesInWorkflow: bit	
PricingErrorCode: int	
EstimatedCloseDate: datetime	
CloseProbability: int	
ActualValue: money	
ActualCloseDate: datetime	
OwningTeam: uniqueidentifier	
OwningUser: uniqueidentifier	
OwningBusinessUnit: uniqueidentifier	
OriginatingLeadId: uniqueidentifier	
CreatedOn: datetime	
IsPrivate: bit	
CreatedBy: uniqueidentifier	
ModifiedOn: datetime	
ModifiedBy: uniqueidentifier	
VersionNumber: timestamp	
StateCode: int	
StatusCode: int	
IsRevenueSystemCalculated: bit	
CampaignId: uniqueidentifier	
PK	
+ PK_OpportunityBase(uniqueidentifier)	

De waarden die relevant zijn uit de OpportunityBase-tabel zijn:

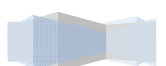
- EstimatedClosedate; dit is de datum waarop de order gaat vallen;
- CloseProbability; de kans dat de order valt;
- EstimatedValue; het bedrag van de order.
- StatusCode; de status van de order.
 - o 4=order gaat niet door
 - o 3=order gaat door
 - o 1=kans is tussen de 0 en 100 %



Wat we willen weten is het volgende. De som van alle orders (SUM EstimatedValue) maal de kans dat de order valt (CloseProbability) en dat voor alle orders binnen een bepaalde maand in het huidige jaar (EstimatedCloseDate). Daarbij mag de order geen status 4 hebben, dit zijn namelijk de orders die niet doorgaan en de status "Lost" hebben. Voorbeeld:

EstimatedCloseDate	CloseProbability	EstimatedValue	StatusCode
2007-04-19	80	25000	1
2007-04-23	60	40000	1
2007-04-12	100	10000	3
2007-04-23	25	12000	4

Om de orderportefeuille voor de maand april te kunnen berekenen vermenigvuldigen we de kolom CloseProbability met de kolom EstimatedValue en delen dit getal door 100. Dit doen we voor alle rijen waarbij de statuscode 1 is en de maand april is. Willen we de orders weten die doorgaan kijken we naar de statuscode die gelijk moet zijn aan 3.

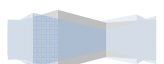


Opstellen feitenmodel

Onderstaande tabel toont alle feiten die we willen gaan meten. De feiten zijn toestanden op een bepaald moment. Nemen we bijvoorbeeld het feit “bruto toegevoegde waarde”, dan bedoelen we in dit geval; de bruto toegevoegde waarde voor een bepaalde businessunit in een bepaalde maand. Dimensies zijn waarden waarin we feiten kunnen indelen. Voor dit project maken we gebruik van slechts drie dimensies; tijd, project en businessunit. De derde kolom toont welke details we opnemen voor de dimensies. Zo nemen we voor de tijdsdimensie slechts het maandgegevens op en voor de businessunit slechts de naam van de businessunit.

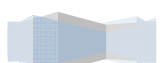
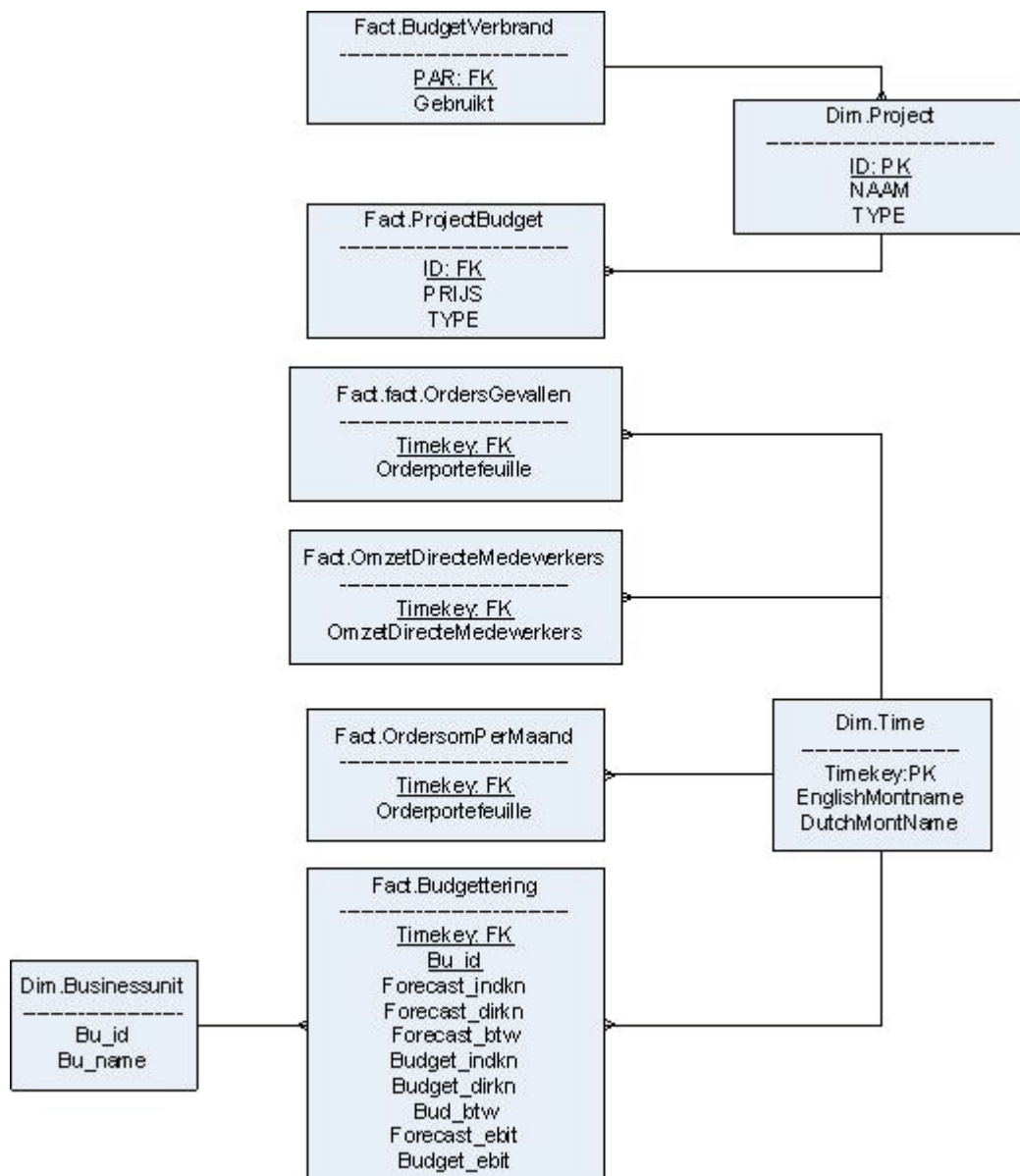
Tot slot is er een kolom opgenomen waarin de brontabel staat waar de feitgegevens te vinden zijn.

Feit	Dimensies	Gegeven	Brontabel
Forecast directe kosten	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Forecast indirecte kosten	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Forecast Bruto toegevoegde waarde	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Budget directe kosten	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Budget indirecte kosten	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Budget Bruto toegevoegde waarde	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Forecast Ebit	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Budget Ebit	Tijd, businessunit	Maand, businessunitnaam	Excelbestand
Verbrand budget	Project	Projectnaam	U_uren en u_structuur
Projectprijs	Project	Projectnaam	U_structuur
Omzet Directe Medewerkers	Tijd	Maand	Excelbestand
Orderportefeuille	Tijd	Maand	OpportunityBase



Opstellen sterschema

Onderstaand sterschema toont het ontwerp van het datawarehouse. Zoals te zien is er slechts een soort relatie, een veel op 1 relatie tussen feitentabel en een dimensietabel. Er is een duidelijk onderscheid aanwezig tussen de feitentabellen en de dimensietabellen door de benamingen van de tabellen.



Surrogaatsleutels

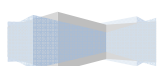
De feiten- en dimensietabellen zijn gekoppeld met behulp van surrogaatsleutels. Dit zijn sleutels die in de originele databronnen (vaak) niet aanwezig zijn en ook geen betekenis hebben. Deze sleutels worden toegevoegd aan iedere rij van een dimensie en worden foreign-keys van de feitentabellen om een koppeling tussen beide tabellen mogelijk te maken. Zo zien we dat er een aantal tabellen zijn gekoppeld met de tijdsdimensie door middel van een timekey. Deze timekey was in de huidige bronnen niet aanwezig. Er werd bijvoorbeeld gebruik gemaakt van een datumveld. Deze zal tijdens het ETL-proces moeten worden omgezet naar een bruikbare surrogaatsleutel om een koppeling mogelijk te maken met de tijdsdimensie. Nu zien we ook het belang terug van een gegevensmodel, we kunnen hier immers in zien welke velden we gaan gebruiken en welke data omgezet moet worden.

De feitentabel Budgettering is verbonden met de businessunitdimensie door middel van de key "Bu_id". Dit is de waarde 1,2 of 3 welke staan voor een van de drie aanwezige businessunits. Deze dimensie zorgt ervoor dat de budgetteringsgegevens kunnen worden gepresenteerd naar businessunit.

Tot slot hebben we de feitentabellen Budgetverbrand en ProjectBudget. Deze bevatten gegevens over de projecten. Om de budgetverbranding te kunnen weergeven, gerangschikt in projecten, is het noodzakelijk om deze te koppelen met de projectdimensie. Hiervoor gebruiken we de sleutels ID en PAR welke gekoppeld zijn met de primaire sleutel "ID" uit de tijdsdimensie. Dit is ook de originele sleutel uit het bronbestand van clockwise.

Granulariteit

Het datawarehouse is ontworpen voor specifieke analysedoelen waardoor er niet veel detail noodzakelijk is. Er zijn geen historische gegevens aanwezig in het datawarehouse. Het voordeel hiervan is dat de database klein blijft en een hoge performance heeft. Het nadeel is dat we geen oude gegevens meer kunnen opvragen en dus geen voorspellingen richting toekomst kunnen doen, echter dit was ook niet de doelstelling van dit project. Kijken we bijvoorbeeld naar de tijdsdimensie dan zien we dat er alleen een maandkolom aanwezig is en geen jaarkolom. Dit heeft als gevolg dat we alleen de gegevens uit het huidige jaar kunnen opvragen aangezien we geen jaartal kunnen kiezen. Hetzelfde geldt voor de dimensie "Project". We slaan slechts de gegevens op die we nodig hebben, dit zijn de attributen NAAM en TYPE, terwijl we ook meer detail hadden kunnen opnemen zoals het aantal medewerkers dat aan een project is gekoppeld of de deadline.



Uitleg tabellen

Fact.Budgetverbrand

In deze tabel staat de aggregatie van de budgetgegevens, namelijk de som van alle uren maal het uurtarief voor een project.

Fact.Projectbudget

In deze tabel vinden we een projectID met een projectbudget. Dit is het budget dat is vastgelegd in de applicatie Clockwise, ook wel bekend als de prijs voor een project. Het ID is afkomstig uit clockwise.

Fact.Ordersgevallen

Hierin staat een timekey en een orderportefeuilleveld. De timekey is een maandnummer. De orderportefeuille is een optelling van alle orders in een bepaalde maand waarbij de statuscode uit de originele tabel op 3 staat. (order gevallen). Hierbij is voor de timekey gekeken naar de EstimatedCloseDate uit het bronbestand.

Fact.OmzetDirecteMedewerkers

Hierin staan de gegevens die nodig zijn om de orderportefeuille tegen af te zetten voor een bepaalde maand. De timekey is het maandnummer uit het budgetbestand.

Fact.OrderSomPerMaand

In het veld Orderportefeuille staat de som van alle orders voor een maand. Dit is de kans * orderbedrag waarvan de statuscode uit originele tabel 1 is (kans tussen de 1-100 procent). Hierbij is voor de timekey gekeken naar de EstimatedCloseDate uit het bronbestand.

Fact.Budgettering

In de budgetteringstabel staan alle gegevens die van belang zijn voor het tonen van financiële data met betrekking tot het budget en de forecast. Zoals te zien is nemen we hierin de volgende gegevens op voor zowel budget als forecast:

Indirecte kosten, directe kosten, btw en ebit. Totaal hebben we dus acht kolommen voor de data en twee voor de koppeling met de twee dimensies tijd en businessunit. De data die hier te vinden is halen we rechtstreeks uit het budgetbestand zonder dat de data omgezet wordt.

Dim.Time

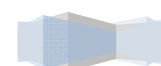
In de tijdsdimensie vinden we de timekey die de waarde 1 t/m 12 kan bevatten. Deze staan voor de maandnummers. Er is geen jaartal opgenomen. Daarnaast vinden we de kolommen EnglishMontname en DutchMontName. Hierin staat voor iedere timekey de vertaling in een zinvolle naam. Dit is nodig om data straks weer te geven in maanden en niet in nummertjes. Eventueel kunnen hier nog andere kolommen aan worden toegevoegd in andere talen. Er zijn op dit moment twee talen opgenomen om de mogelijkheden te tonen.

Dim.Businessunit

In deze tabel vinden we het Bu_id die alleen de waarden 1,2,3 bevat voor de businessunits applications, configurations en detachering voor de kolom bu_name, en is benodigd om de financiële gegevens te groeperen naar businessunit.

Dim.Project

Tot slot is er de dimensietabel Dim.Project. Hierin vinden we de projectnaam, het ID en het type. Het typeveld heeft weinig betekenis maar is gebruikt als controleveld. In de originele tabel kan dit veld de waarde P en R bevatten. We gebruiken alleen de P velden omdat de R velden medewerkers zijn die gekoppeld zijn aan een project.



Opstellen ETL-schema's

Inleiding

Omdat de data zich in meerdere bronnen bevindt zullen we de data moeten overzetten. Uiteindelijk willen we alle data in een datawarehouse hebben staan. Data uitlezen, omzetten en in de nieuwe database plaatsen gebeurt door middel van een ETL-proces (Extract, transform, load). Dit hoofdstuk zal drie schema's tonen zodat duidelijk wordt hoe de data overgezet wordt. Ieder ETL-schema zal worden gekoppeld aan een SSIS-package (SQL Server Integrated Services). Deze packages worden aangemaakt in SQL SERVER 2005 in het onderdeel Integrated Services. We creëren drie afzonderlijke packages. Dit is bewust gedaan omdat niet iedere bron met dezelfde frequentie omgezet hoeft te worden. De uren uit het urensystemen hoeven bijvoorbeeld pas overgezet te worden zodra deze ge-update zijn (eens per week) terwijl de forecastgegevens wel iedere dag ge-update worden.

ETL-schema budgetbestand

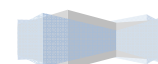
De data die benodigd is om rapportages te maken betreffende de budgetgegevens zijn vastgelegd in een bron, namelijk het budgetbestand (excelsheet). Dit bestand is opgeslagen op de portal van ETTU welke is gemaakt in Microsoft Sharepoint 2007. SSIS (sql server integration services) die verantwoordelijk is voor het ETL proces heeft een opmaak nodig met kolommen (vanuit Excel) om de data over te kunnen zetten. Daarom is het noodzakelijk om een nieuw tabblad binnen het excelbestand te ontwerpen waarin de juiste kolommen zijn gedefinieerd en waarin de juiste gegevens in de juiste kolom staan. Op die manier kunnen we de data op een eenvoudige wijze overzetten naar de SQL database. Het tabblad dat we aanmaken binnen het budgetbestand noemen we "**Warehouse**" en bevat de volgende kolommen:

Tabblad: Warehouse
Maand: integer
Businessunit: integer
Forecast indirecte kosten: money
Forecast directe kosten: money
BTW: money
Gebudgetteerde BTW: money
Gebudgetteerde Indirecte kosten: money
Gebudgetteerde directe kosten: money
EBit Forecast: money
EBit budget: money

Hetzelfde doen we met de gegevens die we nodig hebben voor de KPI die de orderportefeuille afzet tegen het budget. We maken daarvoor het tabblad "**Warehouse-2**" aan en plaatsen daarin twee kolommen. Een kolom bevat het maandnummer die we handmatig vastleggen (1 t/m 12), de tweede kolom bevat verwijzingen naar het tabblad: Consolidatie Budget, velden omzet directe medewerkers.

Tabblad: Warehouse-2
Maandnummer: integer
OmzetDirecteMedewerkers:money

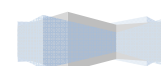
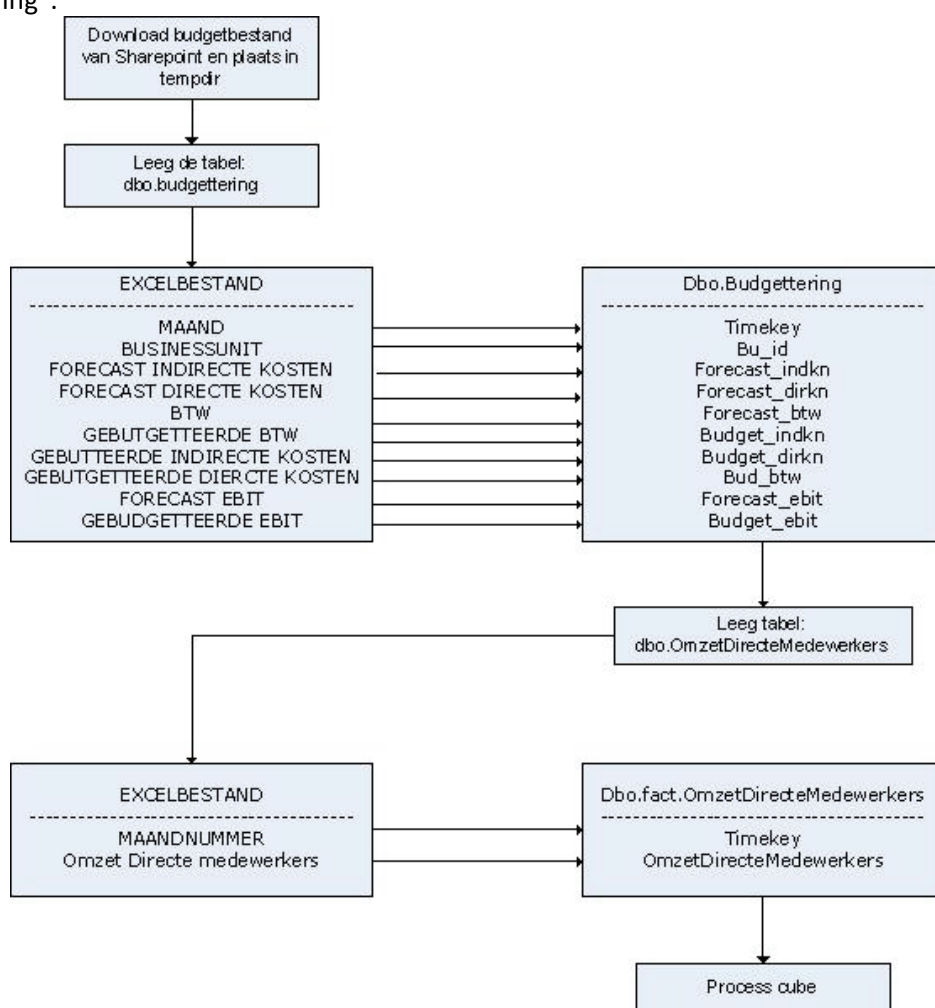
In dit nieuwe tabblad nemen we verwijzingen op naar de gegevens uit de andere tabbladen zodat de juiste waarden in dit tabblad komen te staan. De volgende waarden zijn geen verwijzingen en zijn handmatig vastgelegd: maand, businessunit. De waarde voor het veld businessunit kan



respectievelijk de waarde 1 (Applications), 2 (Configurations) of 3 (Detachering) bevatten. Wanneer er ooit een nieuwe businessunit wordt toegevoegd zal er een nieuwe waarde, bijvoorbeeld 4 worden gekoppeld aan deze unit. De waarde voor de kolom 'maand' leggen we handmatig vast omdat deze waarde altijd een getal tussen de 1 en 12 moet zijn. Kiezen we ervoor om deze over te nemen uit de andere tabbladen (jan, feb) dan moeten we deze data transformeren naar cijfers wat extra tijd kost.

Door dit nieuwe tabblad op te stellen hebben we alle informatie die nodig is voor het kunnen opstellen van de KPI's die gewenst zijn voor het proces bestuur.

Onderstaand ETL schema toont hoe de data wordt ge-extract uit het budgetbestand en wordt overgezet naar het datawarehouse. Zoals te zien is vindt er geen transformatie plaats binnen het proces. De gegevens worden ge-extract vanuit excel en geladen in de nieuwe databasetabel "Budgettering".



Stap 1

Om de data uit Excel te kunnen extraheren moeten we een fysieke locatie opgeven voor dit bestand. Omdat het budgetbestand in Sharepoint 2007 is opgeslagen (in de database) kan er geen directe connectie worden gemaakt met dit bestand. Een oplossing is gevonden door het bestand te downloaden vanuit de database naar een tijdelijke directory op de lokale schijf van de server, namelijk: c:\documents and settings\Administrator\Local Settings\Temp\tempbudget. Deze map kan eventueel veranderd worden naar een zelf te definiëren locatie. Deze methode is gevonden op de pagina: <http://blogs.3sharp.com/Blog/grege/archive/2006/07/12/1534.aspx>.

n.b. dit bestand blijft in deze map staan en wordt iedere keer overgeschreven met de nieuwste versie die in sharepoint is opgeslagen. Verandert de naam van het excelbestand dan moet dit worden aangepast in de variabele "RemoteFileUrl" binnen het SSIS-package.

Stap 2

Omdat we met gegevens werken die betrekking hebben op de toekomst is het niet van belang om oude data mee te nemen in het datawarehouse. Daarom zal er tijdens het ETL-proces een SQL-commando worden meegegeven die de huidige data uit het datawarehouse verwijderd. Dit doen we met het TRUNCATE commando. Deze leegt de tabel "fact.budgettering".

Stap 3

Stap drie is het extract en load proces (Data Flow) waarbij de gegevens uit het budgetbestand worden overgezet naar het datawarehouse. De juiste kolommen worden opgegeven en de locatie waarnaar de gegevens worden gekopieerd wordt opgegeven.

Stap 4

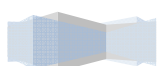
In stap vier wordt de data uit de tabel fact.OmzetDirecteMedewerkers verwijderd. De waarden in deze tabel gebruiken we om de orderportefeuille af te zetten tegen de omzet van de medewerkers in die maand. Omdat we tijdens dit ETL-proces het budget bestand uitlezen nemen we deze actie direct mee.

Stap 5

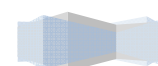
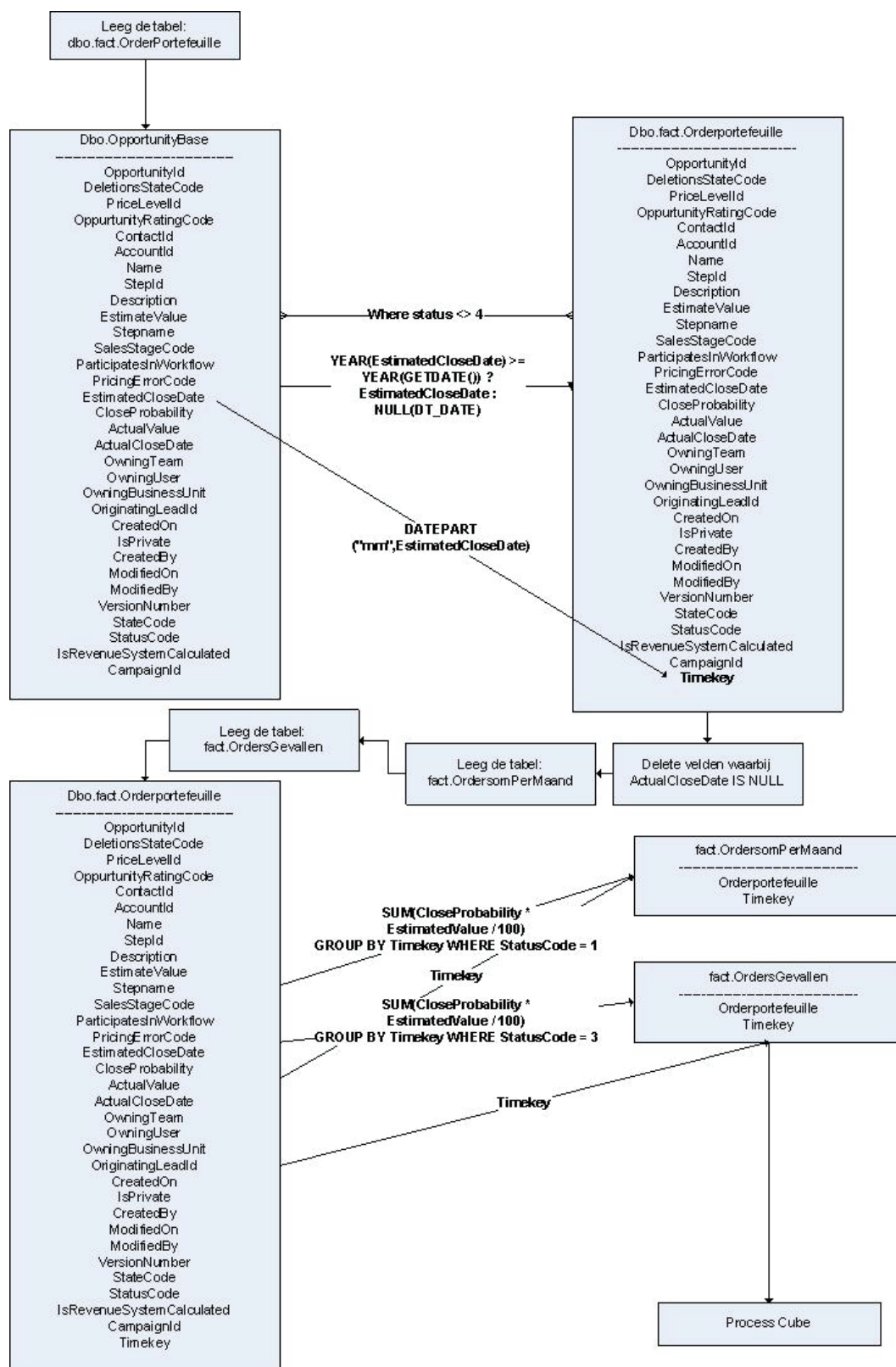
Vervolgens wordt de data uit het tabblad Warehouse-2 welke we eerder hebben aangemaakt overgezet naar het datawarehouse in de tabel fact.OmzetDirecteMedewerkers.

Stap 5

De laatste stap zorgt ervoor dat de cube wordt geprocessed. Dit is nodig om de nieuwe gegevens te laten weergeven.



ETL-schema CRM



Stap 1

Bestaande data wordt uit het datwarehouse verwijderd met het TRUNCATE-commando. Dit doen we voor de tabel "fact.Orderportfeuille".

Stap 2

We maken connectie met de CRM-database en laden alle gegevens uit de tabel "OppurtunityBase". Voorwaarde voor deze selectie is dat alleen de gegevens worden geselecteerd waarbij de statuscode ongelijk is aan "4". Vier staat namelijk voor een "lost order" en is dus niet relevant om mee te nemen. We kijken alleen naar orders die (mogelijk) doorgaan. In de toekomst zou het relevant kunnen zijn ook de lost orders bij te houden, daarvoor zal dan een nieuwe tabel gemaakt kunnen worden waarbij de orders met status 4 worden bijgehouden.

Stap 3

Daarna volgt een "Derive Column" actie in de dataflow waarbij gekeken wordt welke orders in het huidige jaar vallen. Als dit niet het geval is wordt deze rij wel meegenomen maar plaatst de actie de waarde NULL in het veld "EstimatedCloseDate".

Stap 4,5

De volgende twee stappen legen de tabellen "ordersompermaand" en "ordersgevallen" om plaats te maken voor de nieuwe data.

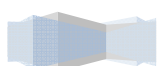
Stap 6 en 7

Nu alleen relevante orders in de eerder gevulde (tijdelijke) tabel "fact.OrderPortefeuille" staan kunnen we de data gaan aggregeren zodat alleen de data overblijft die we nodig hebben. We vermenigvuldigen per order de waarde uit de kolom "EstimatedValue" met het slagingspercentage van de order (de kolom "CloseProbability") en delen dit door 100. Dan tellen we alle orders voor die maand bij elkaar op zodat we de totale orderportefeuille voor die maand hebben. Deze waarden plaatsen we in de tabel "OrderSomPerMaand", samen met de Timekey om de orderportefeuille straks per maand te kunnen tonen.

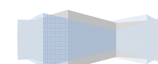
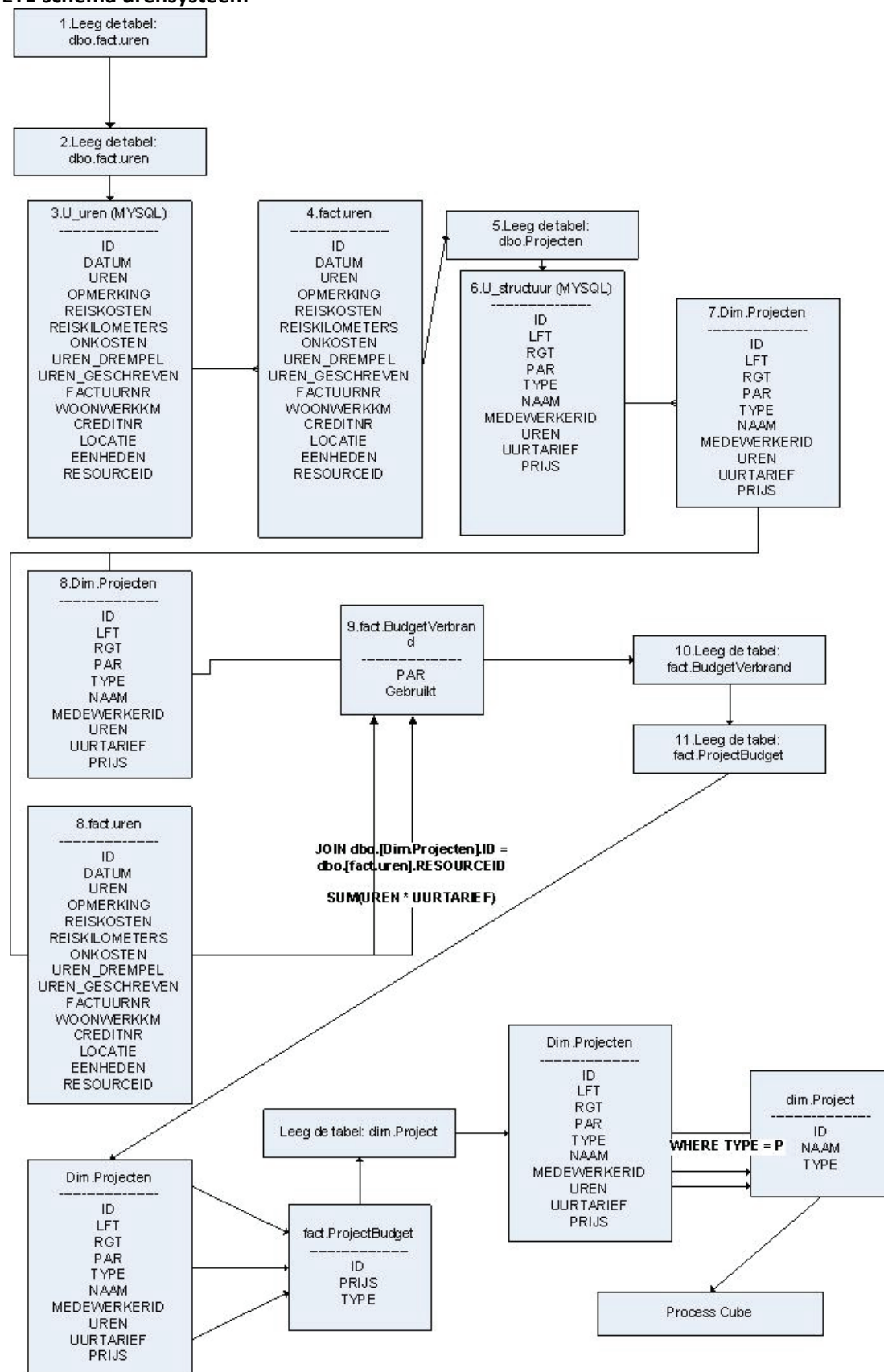
Stap zeven is hetzelfde als stap 6, alleen laden we nu alleen de orders die status 3 hebben (order gevallen) en zetten deze in de tabel "fact.Ordersgevallen".

Stap 8

De laatste stap zorgt ervoor dat de cube wordt geprocessed. Dit is nodig om de nieuwe gegevens te laten weergeven.



ETL-schema urensysteem



Stap 1

In de eerste stap wordt de urentabel opgeschoond zodat er geen dubbele gegevens in voorkomen.

Stap 2

Vervolgens wordt er connectie gemaakt met Clockwise door middel van een ODBC-verbinding. Deze driver is gratis te downloaden (<http://dev.mysql.com/downloads/connector/odbc/3.51.html>). Alle gegevens uit de tabel u_uren wordt geselecteerd en verplaatst naar het datawarehouse in de tabel fact.uren. Het is nodig om eerst alle gegevens over te zetten omdat we op een MYSQL database niet alle commando's kunnen gebruiken die we later nodig hebben.

Stap 3

Hier wordt de projectentabel geleegd.

Stap 4

Er wordt weer connectie gemaakt met clockwise en de velden die nodig zijn worden ge-extract uit de database en getransfereerd naar het datawarehouse in de tabel Dim.Projecten.

Stap 5

De tabel fact.Urenverbrand wordt geleegd.

Stap 6

Hier worden door middel van een SQL commando gegevens uit de tabellen Dim.Projecten en fact.uren geladen en overgezet naar het een andere tabel. Dit commando zorgt ervoor dat er slechts twee kolommen worden overgezet:

- PAR (het projectID)
- Gebruikt (het gebruikte budget, dit aantal uren * uurtarief van alle medewerkers bij elkaar opgeteld)

Stap 7

De tabel fact.Projectbudget wordt leeggemaakt. In deze tabel staan de projectnummers en de budgetten van de projecten.

Stap 8

De tabel fact.Projectbudget wordt opnieuw gevuld met nieuwe waarden voor onderstaande kolommen.

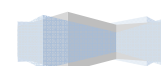
- ID (koppeling met Dim.Project)
- PRIJS (dit is het budget)
- TYPE (om straks te filteren op projecten)

Stap 9

Dim.Project wordt geleegd. Deze tabel is een dimensietabel die informatie bevat over de projecten en een projectnaam koppelt aan een projectnummer. Eventueel kan deze tabel in de toekomst extra attributen gaan bevatten om bijvoorbeeld te filteren op deadline of andere waarden die in de originele u_structuur van clockwise staan.

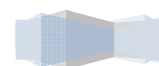
Stap 10

Waarden ID, NAAM en TYPE van het TYPE=P worden uit de Dim.Projecten tabel gehaald en vervolgens geplaatst in de tabel Dim.Project.



Stap 11

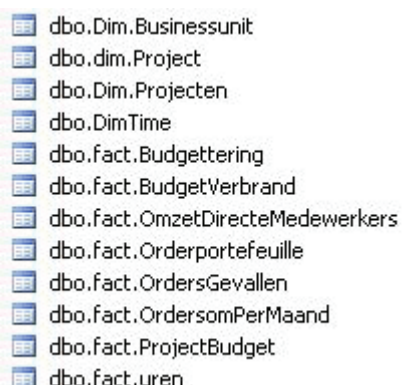
De laatste stap zorgt ervoor dat de cube wordt geprocessed. Dit is nodig om de nieuwe gegevens te laten weergeven.



SQL-database

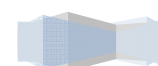
Het datawarehouse zal worden gemaakt in een Microsoft SQL 2005 database welke op een server geïnstalleerd zal worden. De database zal bestaan uit alle tabellen die nodig zijn voor het vormen van de rapportages en KPI's en enkele tabellen die dienen als opslag van tijdelijke data.

Onderstaand figuur toont alle tabellen die in de ettu_dwh database aanwezig zijn. Vervolgens zal daaronder in de tabel worden uitgelegd waar de tabellen voor dienen.



dbo.Dim.Businessunit
dbo.dim.Project
dbo.Dim.Projecten
dbo.DimTime
dbo.fact.Budgettering
dbo.fact.BudgetVerbrand
dbo.fact.OmzetDirecteMedewerkers
dbo.fact.Orderportefeuille
dbo.fact.OrdersGevallen
dbo.fact.OrdersomPerMaand
dbo.fact.ProjectBudget
dbo.fact.uren

Tabel	Functie
Dim.Businessunit	Bevat alle businessunits van ETTU
Dim.Project	Bevat alle projecten van ETTU
Dim.Projecten	Tijdelijke tabel met alle projecten en resources (mensen)
DimTime	Bevat alle maandnummers met de maandnamen
Fact.Budgettering	Alle forecast en budgetgegevens voor het huidige jaar
Fact.BudgetVerbrand	Een overzicht van alle projecten en het gebruikte budget
Fact.OmzetDirecteMedewerkers	Een overzicht die afkomstig is uit het budgetbestand vd omzet
Fact.Orderportefeuille	Bevat alle orders van het huidige jaar en dient als tijdelijke tabel. Hierin staan de orders met statuscode 1,3 en 4.
Fact.OrdersGevallen	Alle orders per maandnummer van het huidige jaar die een statuscode 3 hadden in de originele database van CRM.
Fact.OrdersomPerMaand	Alle orders per maandnummer van het huidige jaar waarvan het slagingspercentage tussen de 0 en 100 procent ligt en de status 1 heeft (nog lopend).
Fact.ProjectBudget	Projectnummers en hun prijs (budget)
Fact.uren	Tijdelijke opslag van geschreven uren die afkomstig zijn het clockwise urensysteem.

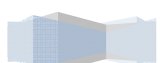
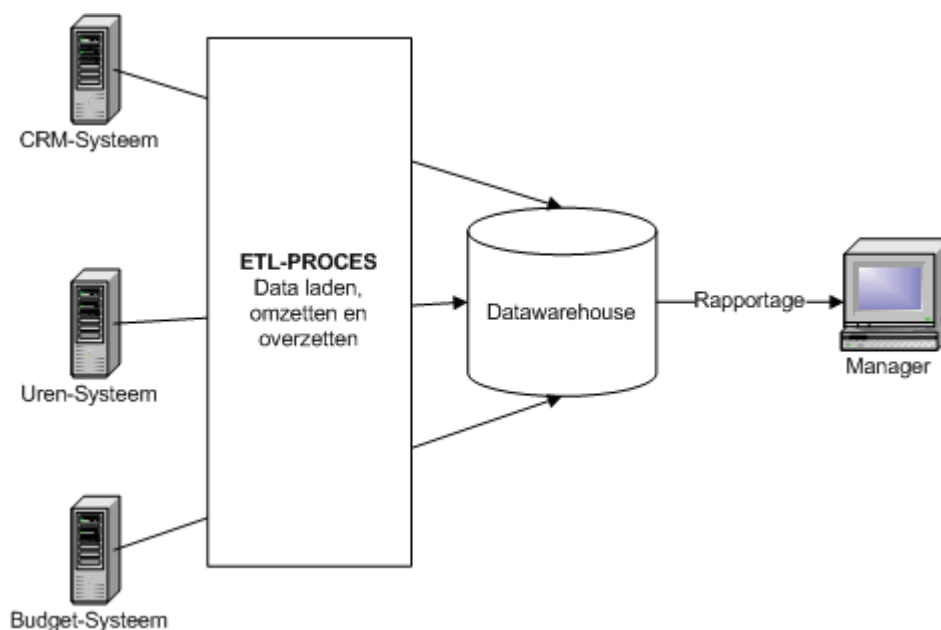


Bijlage E: Implementation Report

Inleiding

Dit document behandelt belangrijke aspecten waar op gelet moet worden bij het overzetten van de business Intelligence oplossing in de productieomgeving.

Eenzijds zijn er de bronsystemen die informatie bevatten over bepaalde aspecten van de business. Zo is er een CRM-systeem, een systeem voor het bijhouden van urenverantwoording en een budgetteringssysteem. De gegevens zijn dus verspreid over meerdere gegevensbronnen. Om de juiste managementinformatie naar voren te krijgen moesten deze gegevens geïntegreerd worden naar een centrale database, het datawarehouse. Het overzetten van de data gebeurt met behulp van zogeheten ETL-packages (Extract, Transform, Load). Het datawarehouse is een OLAP-cube. Dat betekent dat er met feiten en dimensies wordt gewerkt om gegevens overzichtelijk weer te kunnen geven. Uiteindelijk is de manager in staat de juiste data op ieder moment van de dag te raadplegen via Sharepoint 2007. Hieronder is een schematische weergave te zien van de werking van het systeem.



Meegeleverde bestanden

Dit project bestaat uit een aantal bestanden en mappen.

Budget en Forecast 2007 - 2007 versie.xlsx

Dit is de meest recente versie van het budgetbestand waarbij de nieuwe tabbladen aanwezig zijn.

Cockpit.xlsx

Bestand met draaitabellen en grafieken die wordt weergegeven in Sharepoint door middel van Excel Services.

Database_SQL

Een dump van het datawarehouse die in SQL Server moet worden geïmporteerd.

Ettu Cube 2007.odc

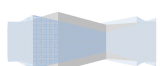
Een verbindingsbestand die Excel gebruikt om connectie te maken met de Cube. Dit bestand moet in Sharepoint worden gezet in de Data Connection map en ook worden aangepast met de juiste instellingen.

Map: Ettu 6-4-2007

In deze map staan alle projectfiles van de Cube die gebouw is in MS Analysis Services

Map: ETTU ETL PROJECT

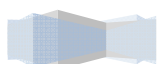
In deze map staan de ETL-packages die gebouw zijn in MS Integration Services.



Microsoft SQL Database

Alvorens te starten moet er bekeken worden of Service Pack 2 voor SQL Server is geïnstalleerd op de server. De data die nodig is voor het weergeven van de juiste data in het managementdashboard ligt opgeslagen in een Microsoft SQL Server. Deze SQL database vormt het uitgangspunt van het project. Onjuiste data in deze database resulteert in een verkeerde weergave van managementinformatie. Bij het implementeren kan de database worden overgezet naar de nieuwe omgeving.

Meer informatie over de tabellen in de SQL-database is te vinden in het designrapport.



ETL-processen

De ETL-packages zorgen ervoor dat de data uit de bronsystemen wordt gelezen, getransformeerd en wordt overgezet naar het datawarehouse. Totaal zijn er drie ETL-packages gemaakt. Voor ieder bronsysteem is er een ETL-package opgesteld. Onderstaande drie ETL-packages zijn ontwikkeld:

- ETL-BUDGET.dtsx
- ETL-CLOCKWISE.dtsx
- ETL-CRM.dtsx

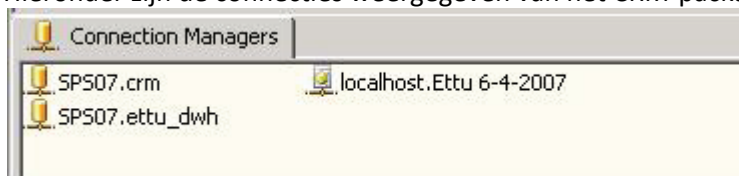
De ETL-packages worden standaard opgeslagen in de projectenmap op de schijf van de server.

Bij het overzetten naar de live-omgeving moet er gelet worden op onderstaande zaken.

Connection Managers

Alvorens een ETL-package wordt uitgevoerd zal er een connectie worden gelegd met het bronsysteem om de data uit te kunnen lezen. Daarnaast is er een tweede connectie aanwezig met het datawarehouse waar de data in geplaatst wordt. De laatste connectie is die met de OLAP-cube om deze te kunnen uitvoeren. Ieder ETL-package heeft dus drie verbindingen nodig.

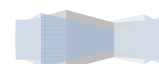
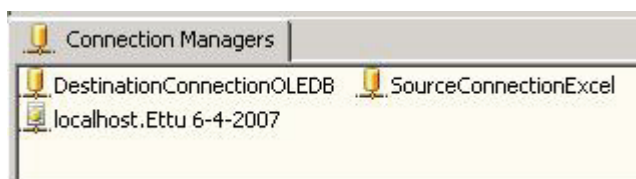
Hieronder zijn de connecties weergegeven van het CRM-package.

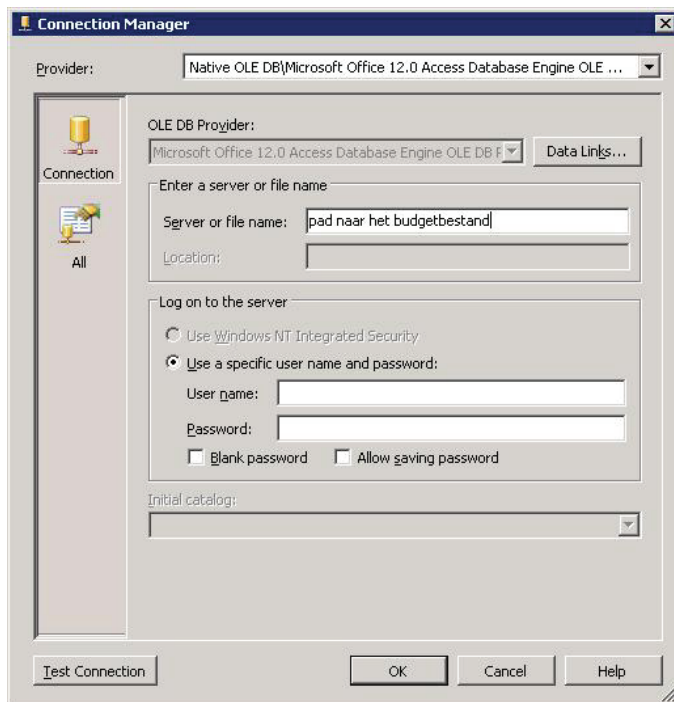


Het ETL-proces van het Budgetbestand bestaat uit drie connecties. SourceConnectionExcel connect met een office 2007 file. Destination is het ETTU warehouse en localhost is nodig om de cube te processen.

URL naar een voorbeeldconnectie voor Office 2007 bestanden:

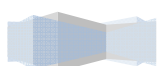
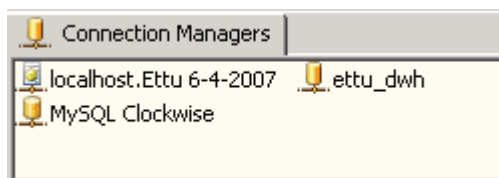
[http://ssis.wik.is/Data_Sources/Microsoft_Office_\(2007\)](http://ssis.wik.is/Data_Sources/Microsoft_Office_(2007))





Tot slot is er het ETL-package dat verantwoordelijk is voor het overzetten van de data vanuit Clockwise. Deze bevat ook drie connections. Let op: clockwise heeft een ODBC-verbinding nodig. De driver hiervoor is gratis te downloaden op onderstaande website:

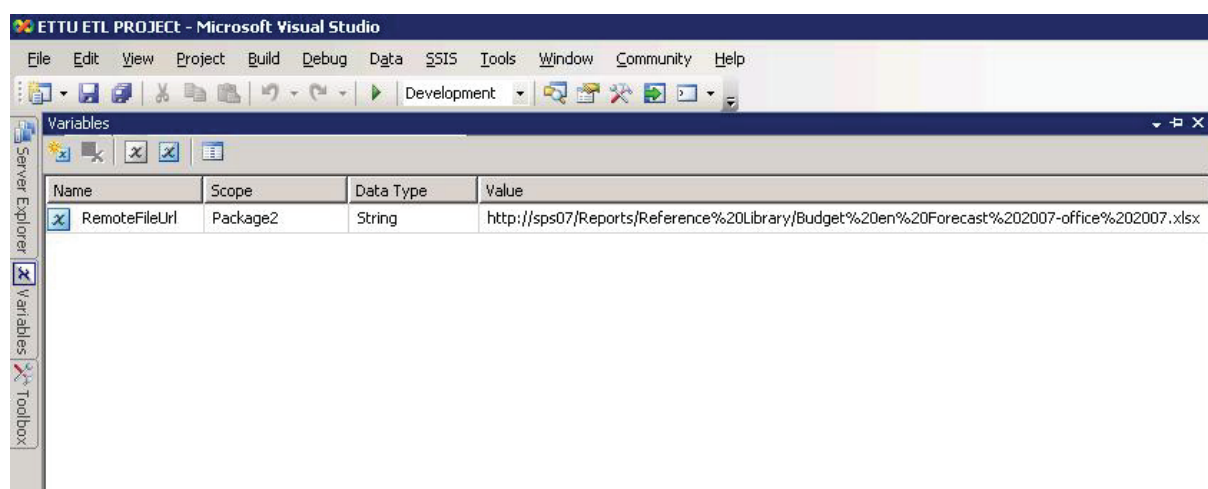
<http://dev.mysql.com/downloads/connector/odbc/3.51.html>



Variabelen

Het ETL-package dat verantwoordelijk is voor het overplaatsen van de financiële data uit het budgetbestand maakt gebruik van een variabele. In deze variabele kan de locatie van het Excelbestand worden ingevoerd wanneer deze is opgeslagen in een sharepoint omgeving. Het scherm met variabelen kan worden opgeroepen vanuit de linkerkant van het scherm binnen Integration Services. In het veld "Value" moet de volledige URL worden ingevoerd van de locatie waar het Excelbestand zich bevindt.

N.B. wanneer in 2008 de naam van het budgetbestand wijzigt is het noodzakelijk om deze variabele aan te passen, anders kan het bestand niet worden gevonden.

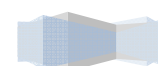


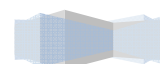
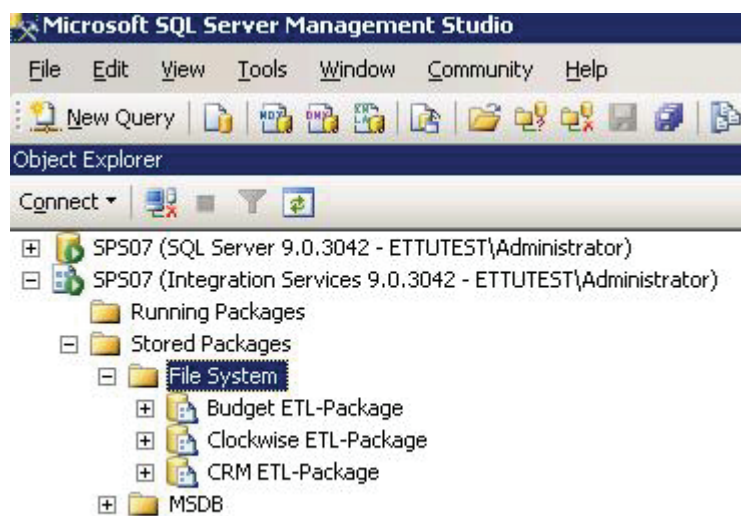
Alvorens het excelbestand kan worden uitgelezen wordt deze gedownload naar een tijdelijke locatie op de server. Om het pad aan te passen van de tempdir waarin het bestand geplaatst moet worden zal er in de eerste workflow "download budgetbestand" de scriptcode gewijzigd moeten worden. De volgende regel moet hiervoor aangepast worden:

```
'set up file paths
localFolder = My.Computer.FileSystem.SpecialDirectories.Temp
' localFolder = System.Environment.GetFolderPath(System.Environment.SpecialFolder.LocalApplicationData)
localFileName = fileUrl.Segments(fileUrl.Segments.Length - 1)
fileDestination = localFolder + "\" + localFileName
```

ETL-Packages in database plaatsen

Om de packages straks te kunnen plaatsen in een schedule (om deze uit te laten voeren) is het van belang dat deze eerst in de database worden geplaatst. Door te connecten via de SQL Server Management Studio met Integration Services kan dit worden gerealiseerd. Ieder package wordt daarna toegevoegd aan de map "Stored Packages" in de submap "File System".

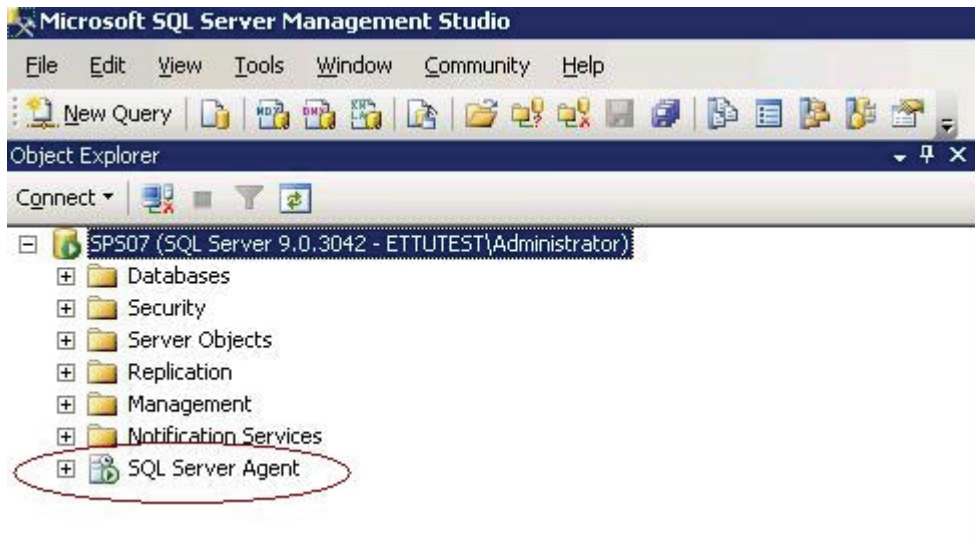




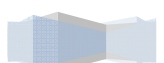
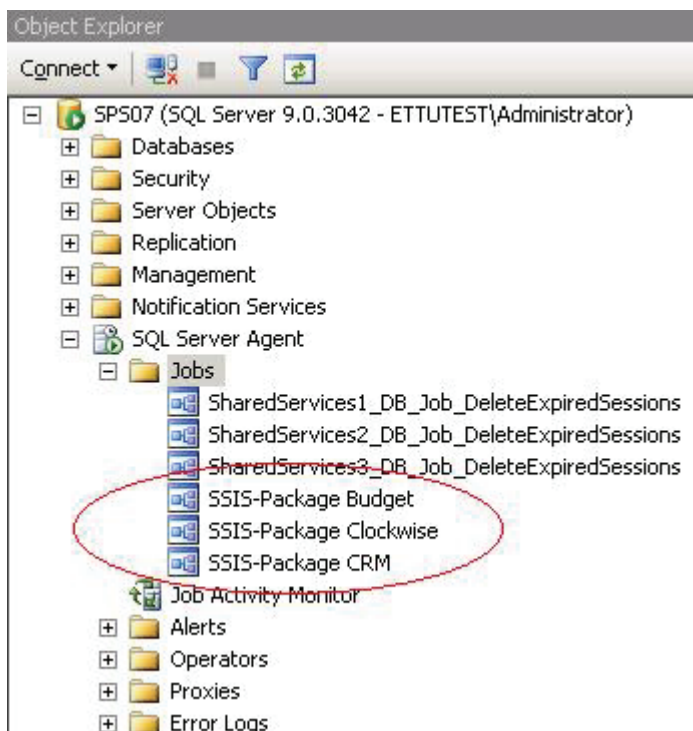
Job-execution

Om ervoor te zorgen dat de ETL-packages automatisch uitgevoerd worden op een vooraf gedefinieerd tijdstip is het noodzakelijk om deze te scheduleren. Dit kan verzorgd worden door de SQL-agent. Deze moet geactiveerd worden vanuit het services menu in Windows.

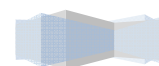
Wanneer de SQL-agent actief is kan deze worden aangepast vanuit SQL Server Management Studio.



Vervolgens moeten er drie taken worden aangemaakt voor de uitvoer van ieder ETL-package. In de taak wordt aangegeven op welk tijdstip het package moet worden uitgevoerd.



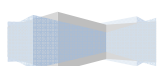
De data uit het budgetbestand zal wellicht iedere dag verversst moeten worden terwijl de data uit clockwise slechts een maal per week hoeft te worden verversst. Verder moet er op worden gelet dat er voldoende rechten zijn om de packages uit te voeren.



Budgetbestand

De aanpassing die aan het budgetbestand nodig zijn vereisen speciale aandacht. Omdat de data binnen het budgetsysteem ongestructureerd is opgeslagen moest deze data geherscht worden naar een formaat dat uitleesbaar is door het ETL-package. Er zijn daarom twee tabbladen aangemaakt waarop de data gestructureerd is vastgelegd. Zonder deze twee tabbladen kan het ETL-package de data niet overzetten. Het is dus van belang deze twee tabbladen toe te voegen aan het budgetsysteem. Eventueel kan ervoor worden gekozen deze tabbladen onzichtbaar te maken zodat ze niet worden weergegeven.

De twee tabbladen die moeten worden toegevoegd zijn te vinden in het budgetbestand dat in de testomgeving te vinden is. Deze twee kunnen worden gekopieerd naar het bestand dat in de operationele omgeving aanwezig is.



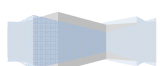
OLAP-Cube

De manier waarop de data kan worden gepresenteerd is vastgelegd in een OLAP-Cube. Hierin staan alle feiten en dimensie gedefinieerd, evenals enkele KPI's. De CUBE is gebouwd in Microsoft Analysis Services in het project "Ettu 6-4-2007".

Zaken die in de cube zijn vastgelegd:

- Feiten;
- Dimensies;
- Berekeningen;
- KPI's.

Het is van belang dit project over te zetten naar de nieuwe omgeving om ervoor te zorgen dat de cube kan worden uitgevoerd. Een verkeerde aanpassing in de cube kan ervoor zorgen dat het hele systeem niet meer functioneert.



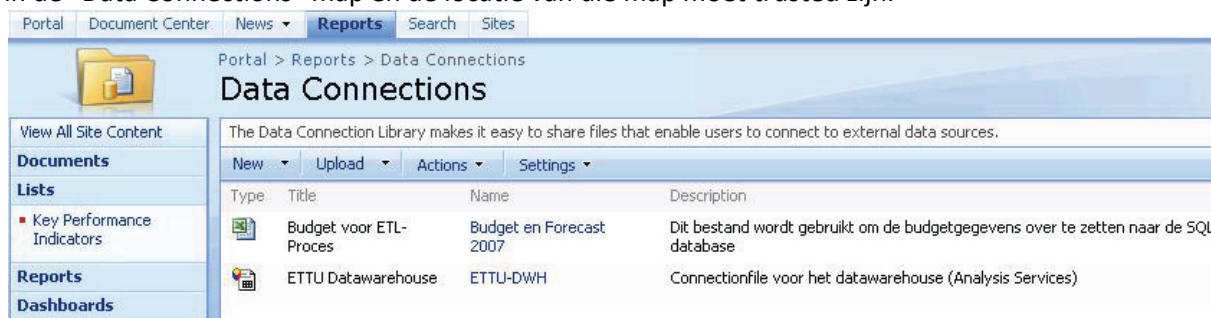
Microsoft Sharepoint 2007



De managementinformatie wordt uiteindelijk weergegeven in een dashboard, ook wel bekend als de cockpit. De data die gepresenteerd wordt is afkomstig uit de OLAP-Cube, vervolgens wordt deze informatie via Excel weergegeven in sharepoint 2007. Daarvoor wordt het bestand cockpit.xlsx gebruikt die is opgeslagen in de reference library van de cockpit.

Excel 2007

De nieuwe versie van Excel biedt mogelijkheden als het gaat om business Intelligence. Zo is het mogelijk KPI's te definiëren en te werken met gekleurde velden. Om de managementinformatie te tonen in Sharepoint wordt er gebruik gemaakt van de Excel services webpart.

Om de data die in de cube ligt opgeslagen weer te geven in Excel 2007 moet er een connectie worden gemaakt met de cube. Dit wordt verzorgd door een vooraf gedefinieerde .odc file. In deze file is de connectie-informatie te vinden. Wanneer de cube wordt verplaatst naar een nieuwe server moet dit bestand dus logischerwijs worden aangepast. Het .odc bestand is opgeslagen in sharepoint in de "Data Connections" map en de locatie van die map moet trusted zijn.



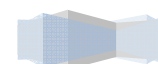
Type	Title	Name	Description
	Budget voor ETL-Proces	Budget en Forecast 2007	Dit bestand wordt gebruikt om de budgetgegevens over te zetten naar de SQL database
	ETTU Datawarehouse	ETTU-DWH	Connectionfile voor het datawarehouse (Analysis Services)

Voor dit project is er een Excel 2007 bestand aangemaakt die connecties bevat naar de OLAP-cube met behulp van bovenstaande connectiefile. In dit cockpit.xlsx bestand zijn de rapportages te vinden die worden weergegeven in het dashboard. Er wordt gebruikt gemaakt van draaitabellen (pivot-tables). Het Excelbestand is opgeslagen in de map "Reference Library" op de cockpitsite van sharepoint.

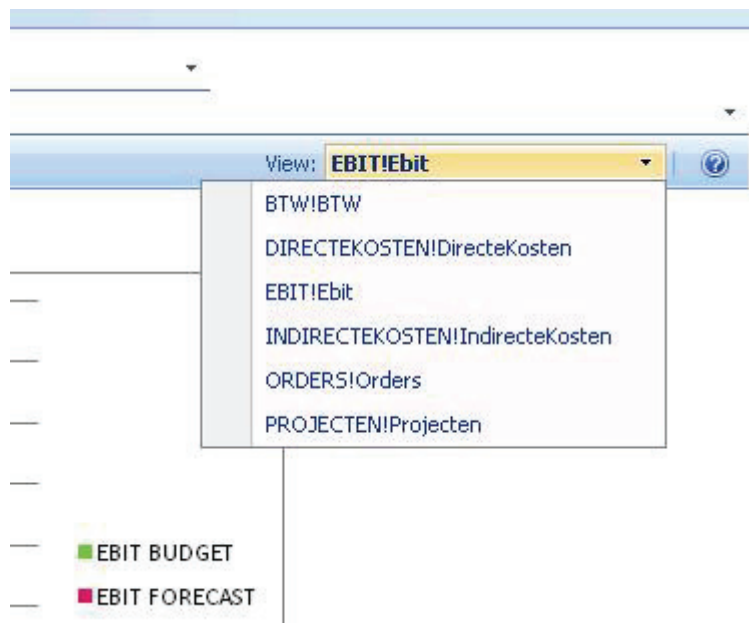


Type	Name
	Budget en Forecast 2007-office 2007
	cockpit

In dit Excelbestand staan diverse tabbladen met data. Ieder tabblad maakt gebruik van een named space. Hierin is een bepaald celbereik gedefinieerd en gekoppeld aan een naam. Dit is nodig om niet de gehele excelsheet in sharepoint te tonen maar enkel dat stukje dat vooraf is gedefinieerd. In de



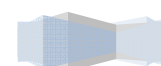
excelsheet kunnen de standaard waarden worden aangepast. Dit zijn de waarden die worden getoond in het managementdashboard. Zoals het Excelbestand wordt opgeslagen in sharepoint, zo worden ook de gegevens getoond in het dashboard. Geven we hierin een aantal projecten aan die getoond worden met behulp van de filter dan verschijnen deze ook standaard in het dashboard. Onderstaand figuur toont welke Named-spaces zijn gedefinieerd in Excel en vervolgens zijn gepubliceerd.

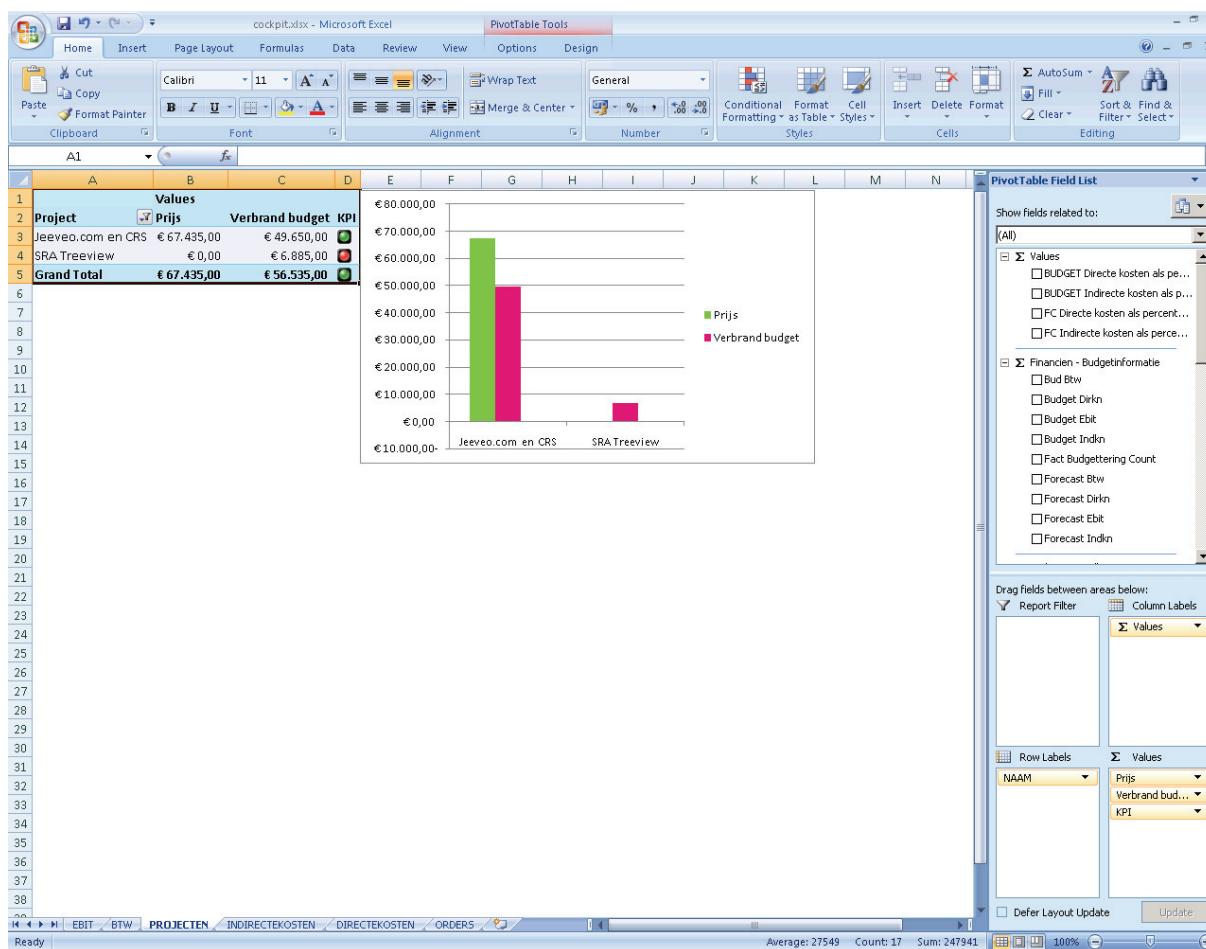


Bij het publiceren van het Excelbestand naar sharepoint 2007 kan er worden gekozen welke named-spaces meegenomen moeten worden. Standaard staan deze al goed ingesteld. Eventueel kan dit aangepast worden.

Verder is er ingesteld dat de data uit de excelsheet automatisch moet worden verversd wanneer deze wordt geopend. Dit zorgt ervoor dat de nieuwste data automatisch op het scherm wordt getoond. Deze optie is te vinden in het tabblad "Data", Refresh All, Connection Properties. Wanneer dit vinkje niet aanstaat zal de data niet automatisch verversd worden.

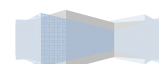
Onderstaand figuur toont het tabblad "PROJECTEN". In dit tabblad is een pivot-table te vinden die verbinding maakt met de OLAP-cube. Daarnaast is een grafiek te vinden van de data die in de pivot-table wordt getoond.



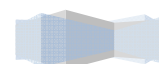
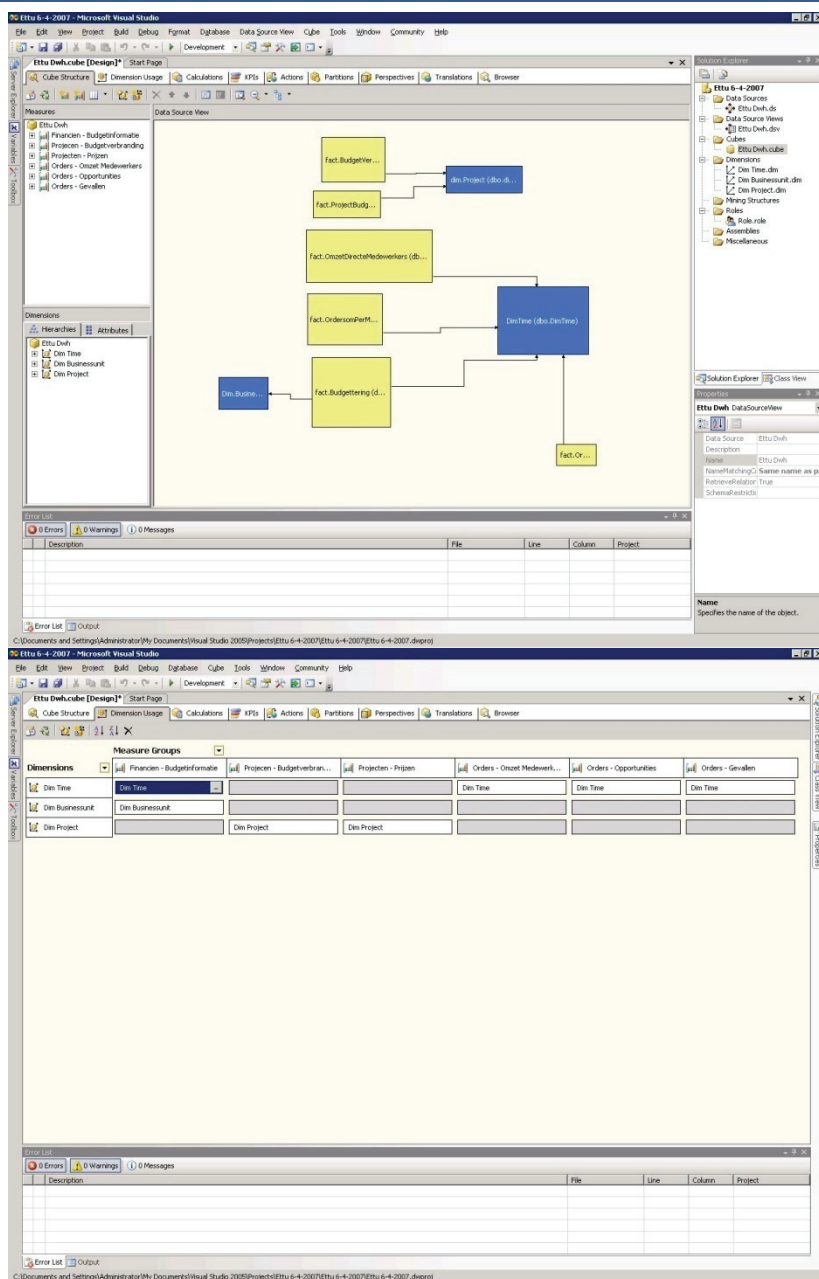


Voor een juiste werking tussen Excel services en sharepoint moet Excel Services in “Delegation Mode” worden gezet.

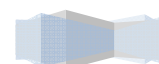
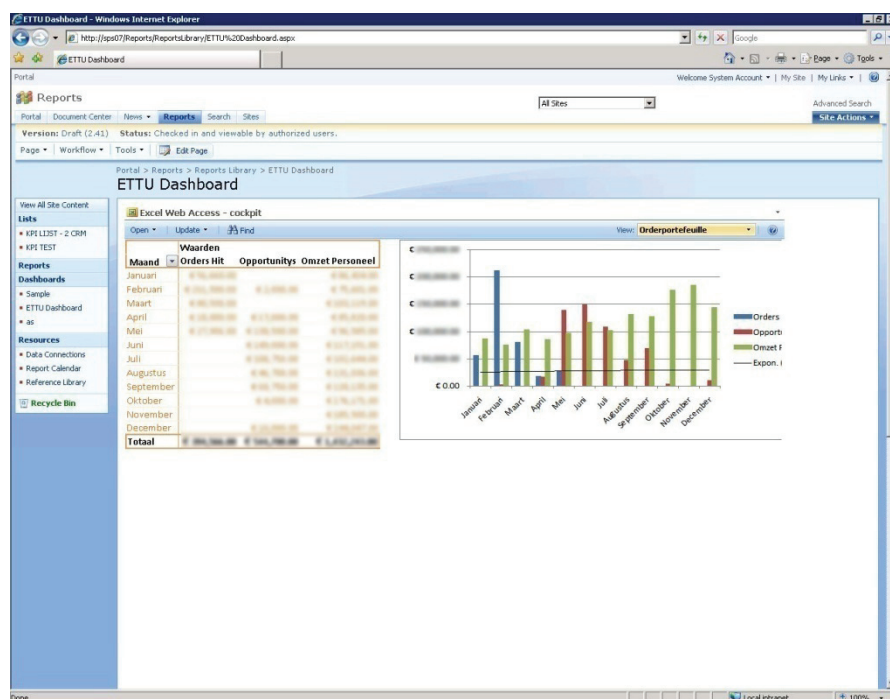
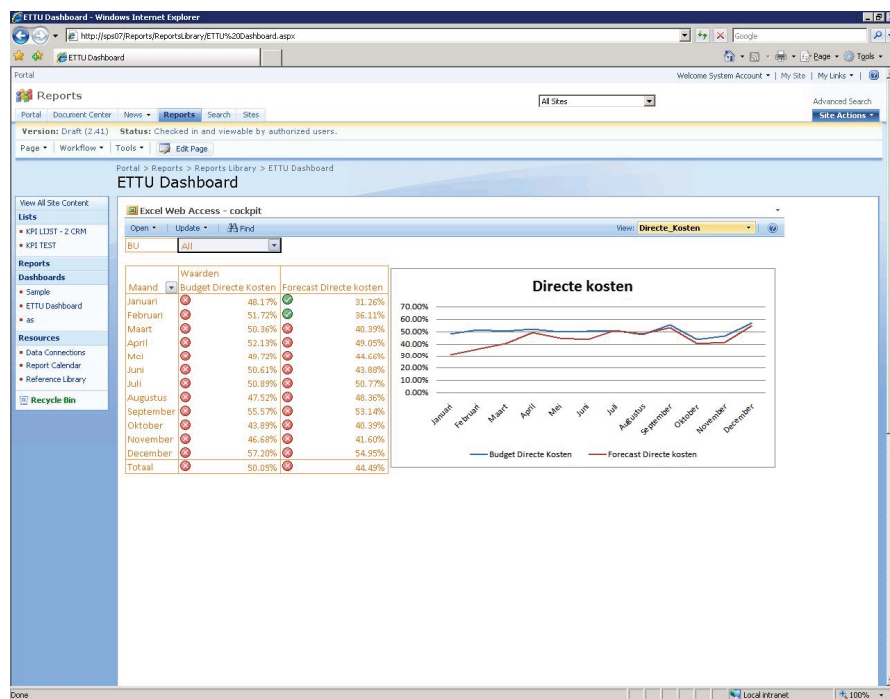
Linkje voor uitleg: <http://blogs.msdn.com/gunterstaes/archive/2007/03/16/data-refresh-failed-in-excel-services.aspx>



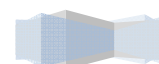
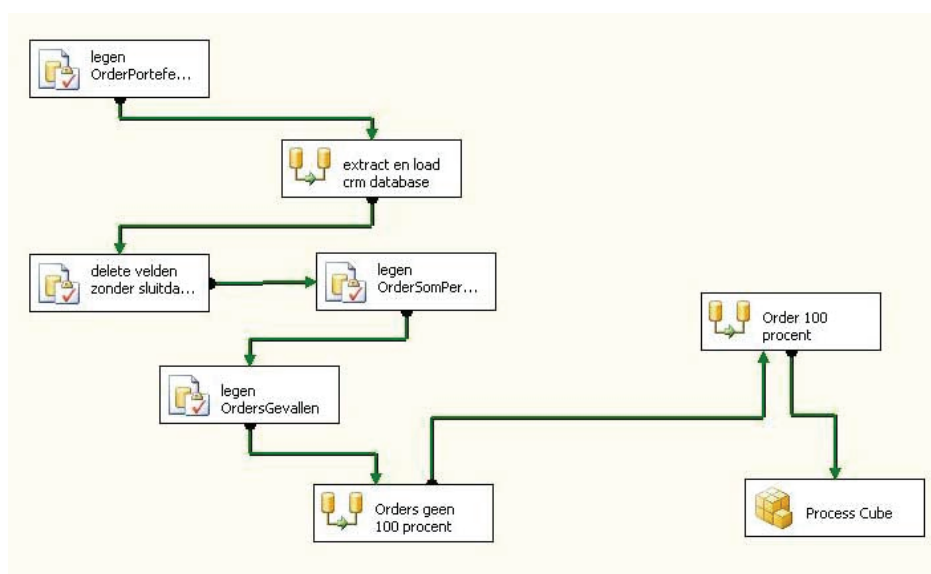
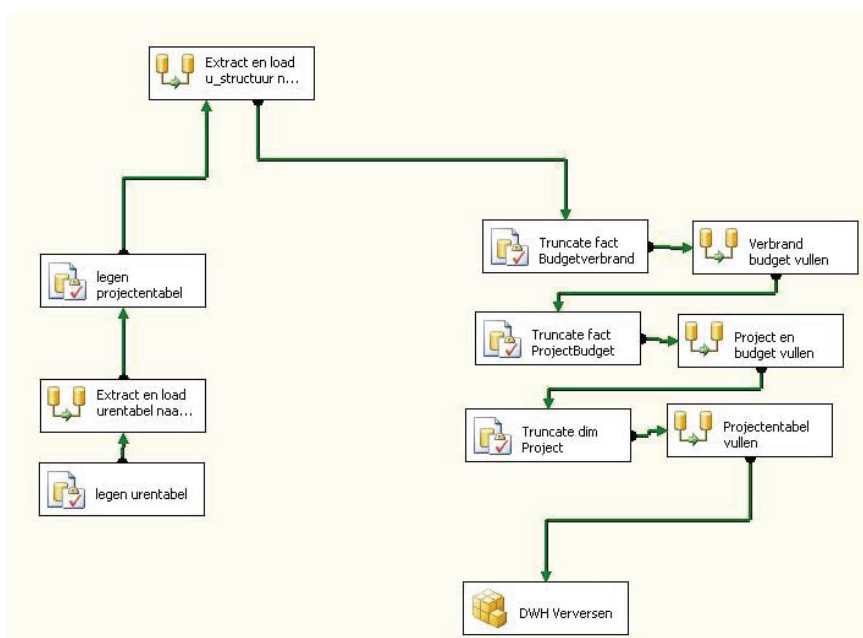
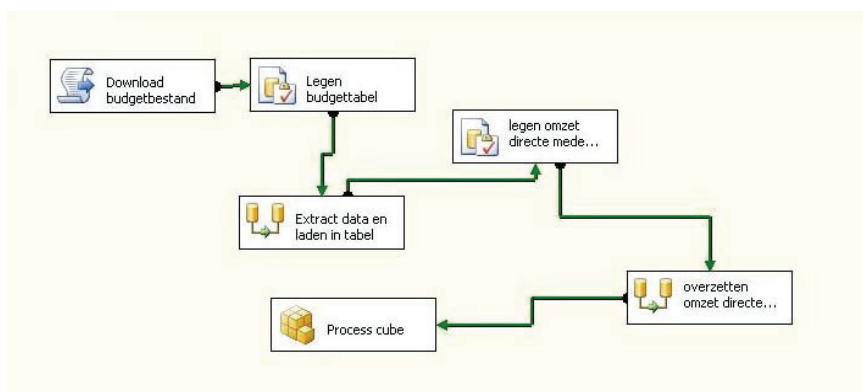
Bijlage F: Screenshots datawarehouse



Bijlage G: Screenshots Managementdashboard



Bijlage H: Screenshots ETL-packages



Bijlage I: Figurenlijst

Figuur 1 Business intelligence bij ETTU.....	10
Figuur 2 Plan van aanpak	12
Figuur 3 Projectrisico's	17
Figuur 4 Fase 1: justification	18
Figuur 5 Fase 2: planning	20
Figuur 6 Fase 3: business analysis	23
Figuur 7 Agenda	25
Figuur 8 Notulen vergadering	26
Figuur 9 Prioriteiten informatiebehoefte	27
Figuur 13 Overzicht KPI's	28
Figuur 14 Uitgewerkte KPI	29
Figuur 10 Opportunities in CRM	31
Figuur 11 Projecten in het urensysteem.....	33
Figuur 15 Datakwaliteit.....	35
Figuur 12 Business Requirements Report.....	36
Figuur 16 Fase 4: Design	38
Figuur 17 Gegevensmodel budgetsysteem.....	39
Figuur 18 Gegevensmodel urensysteem	40
Figuur 19 Gegevensmodel CRM.....	40
Figuur 20 Sterschema datawarehouse	43
Figuur 21 Sterschema datawarehouse uiteindelijke versie	44
Figuur 22 Werking ETL-proces	46
Figuur 23 Bronnenmodel ETTU	48
Figuur 24 ETL-schema budgetsysteem	51
Figuur 25 ODBC driver.....	52
Figuur 26 Timekey toevoegen CRM	53
Figuur 27 Fase 5: Construction	54
Figuur 28 Analysis Services Tutorial.....	56
Figuur 29 meetwaarden en dimensie tijd.....	58
Figuur 30 meetwaarden en dimensies tijd en businessunit	59
Figuur 31 Excel 2007 draaitabel met KPI's	60
Figuur 32 Excel 2007 grafiek	61
Figuur 33 Dashboard in Sharepoint 2007	61
Figuur 34 Fase 6: deployment.....	63

