

# Het Metadata Frame in Actie

## *Ontrafeling van de Babylonische spraakverwarring*

Peter Alons (Dr. P.W.F. Alons)  
Atos Origin/Business Intelligence-CRM



Zeven jaar geleden schreef ik twee artikelen over metadata management in DB/M [1, 2]. Daarin beschreef ik een methode genaamd het Atos Origin Metadata Frame. Deze is gericht op volledige benutting van het conceptuele niveau van het informatieaspect. Sindsdien is onze methode buitengewoon heilzaam gebleken in verschillende branches. In dit artikel beschrijf ik hoe we gebruik hebben gemaakt van de methode bij het Erasmus Medisch Centrum in Rotterdam.

Bij het Erasmus Medisch Centrum in Rotterdam stond de IT-afdeling voor de grote uitdaging om waardevolle informatie te leveren aan de medische staf, in het bijzonder over de zorg in hun Intensive Care afdelingen (IC's). Deze informatie is niet alleen vereist voor rapportage aan de Zorginspectie, maar is ook hard nodig voor hun eigen onderzoek gericht op het verbeteren van de kwaliteit van hun zorg. En hoewel ze over uitstekende operationele informatiesystemen beschikten, bleef hun vermogen om daar informatie uit te halen jammerlijk achter. Of zoals de geneesheer-directeur van de IC voor volwassenen, Prof. Dr. Jan Bakker, het uitdrukte: "Ik kom om in de gegevens, maar ik heb *geen informatie!*"

Uit interviews met IC-artsen, verpleegkundigen en onderzoekers bleek, dat zij adequate informatie nodig hadden om diverse redenen:

- Onderzoek voor kwaliteitsverbetering na reorganisatie van de IC's.
- Analyse van diagnoses: reden voor opname, diagnose bij opname en bij ontslag, en de uiteindelijke, gedetailleerde diagnose.
- Het met groot gemak beantwoorden van simpele vragen als:
  - Hoeveel patiënten zijn in een gegeven periode behandeld door welke arts?
  - Welke artsen hebben welke medicatie toegepast in een gegeven periode?
  - Wat is het percentage van bedgebruik?
- En het krijgen van de juiste antwoorden op moeilijker vragen als:
  - Laten we de juiste patiënten toe?
  - Wat is het aantal doses versus dagen bij medicaties voor patiënten die X doses of meer ontvingen?
  - Wat is de doorstroming van Eerste hulp – Operatiekamers – IC?
  - Enzovoort...

Net als veel bedrijven, probeerde onze klant in dit soort informatiebehoefte te voorzien door het opzetten van een Data Warehouse. En daarbij stuitte zij op dezelfde uitdagingen waarmee veel van onze klanten geconfronteerd worden, als zij een Data Warehouse bouwen om waardevolle informatie te leveren aan hun eindklanten: Hoe kom je aan ondubbelzinnige gegevensdefinities? Hoe zet je metadata management effectief op? Hoe krijg je de eindgebruikers afdoende betrokken bij het ontwerp van een Data Warehouse? En hoe draagt een methodische aanpak bij aan het beantwoorden van deze drie vragen? Dat brengt ons terug bij onze Metadata Frame methode, die we vooral ontwikkeld hebben om deze vragen effectief te beantwoorden:

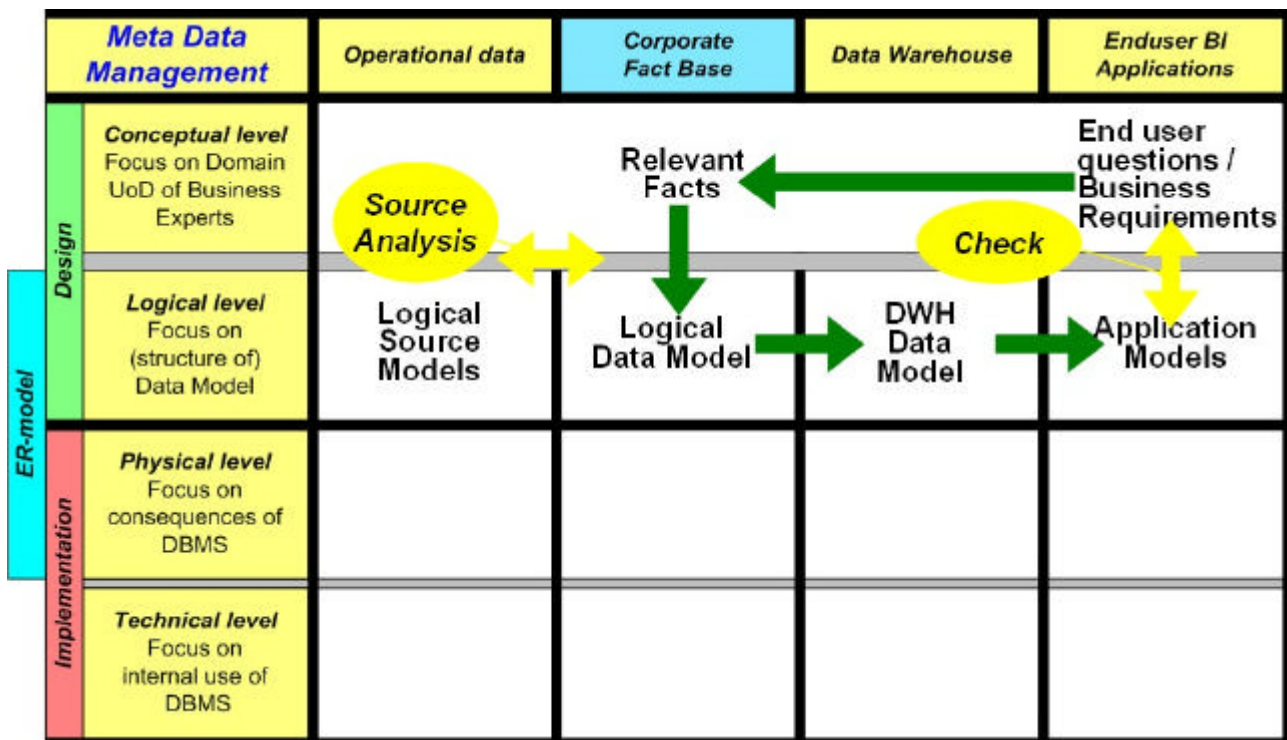
Als we een Data Warehouse bouwen, stellen we ons hoge doelen met hoge verwachtingen, en hebben behoefte aan adequate kennis en documentatie, gepaard aan een evolutionaire realisatie in een complexe en veranderlijke omgeving. Vaak is die omgeving gefragmenteerd, met verschillende codes voor dezelfde dingen en meer dan een operationeel systeem voor hetzelfde doel. Geen wonder dat we in zulke projecten serieus worden bedreigd met een Babylonische spraakverwarring. De standaard IT aanpak om deze problemen het hoofd te bieden is een poging om een bedrijfsbreed 'business model' op te zetten in termen van een Entity-Relationship model. Helaas leidt deze aanpak niet tot ontrafeling van de Babylonische spraakverwarring.

## De ontwikkeling van informatiemodellen

Om een Data Warehouse te bouwen hebben we toegesneden datamodellen nodig. Entity-Relationship Modelling (ERM) is om een aantal redenen geen goed middel om het creatieproces van de vereiste datamodellen aan te sturen. Allereerst zijn de meeste bedrijfsmensen buiten het werkgebied van de IT niet vertrouwd met ERM. Daarom kunnen we niet verwachten, dat communicatie met de business experts in termen van ER-diagrammen doeltreffend kan zijn. Ook is het niet ethisch om met iemand een gesprek op te zetten, waarvan de uitkomst voor hem van vitaal belang is, in een taal die hij niet begrijpt. Op zijn best is deze aanpak doelmatig, maar dan alleen omdat we helemaal niet communiceren met de business experts. Maar dan moeten we niet verbaasd zijn, als de uiteindelijke uitkomst van de bouw van het Data Warehouse inefficiënt en ineffectief blijkt te zijn.

Figuur 1 toont het Metadata Frame zoals het functioneel door ons wordt gebruikt. Dit frame richt zich op het informatieperspectief van informatiesystemen, en verdeelt dit perspectief in vier niveaus:

- Het *Conceptuele* niveau, waar de focus op het domeingebied ligt: het 'Universe of Discourse' van de business experts. Dit is feitelijk het 'eindgebruiker niveau', of iets fraaiër: 'eindklant niveau'.
- Het *Logische* niveau, waar de focus op het datamodel ligt: de structuur van en de relaties tussen alle entiteiten en gegevenselementen. Dit is het 'ER niveau'<sup>1</sup>.
- Het *Fysieke* niveau, waar de focus op het fysieke platform ligt: de consequenties voor het relationele model van de gegevenselementen. Dit omvat het toevoegen van afleidbare informatie, het benoemen van de 'foreign key' elementen, en naamsveranderingen die worden afgedwongen door het DBMS dat wordt gebruikt. Dit is het 'RM niveau' (Relationeel Model).
- Het *Technische* niveau, waar de focus op het internal gebruik van het gekozen DBMS ligt. Dit omvat alle indexes, 'views', en 'partition-keys' die de DBA als geschikt beschouwt. Dit is het 'DBMS niveau'.



Figuur 1: Het Metadata Frame en het functioneel gebruik daarvan in het ontwerp proces.

De bovenste twee niveaus van het frame vormen de *ontwerplaa*g van het data modelleringsproces, de onderste twee de *implementatielaag*. Het logische en fysieke niveau vormen samen de *ER-laag*, zoals ook is te zien in alle ER-tools zelf. Essentieel in het gebruik van het Metadata Frame is, dat we de vier niveaus strikt gescheiden houden: we lossen de problemen op een gegeven niveau op tijdens werk op dat niveau zelf, niet als we aan het werk zijn op een ander niveau.

<sup>1</sup> Bij gebruik van Class Diagrams kan men hier ook *UML*-niveau lezen.

De kolommen in het Metadata Frame worden gevormd door de diverse 'zuilen' van toepassing. In fig. 1 zijn dit zuilen die geschikt zijn voor het bouwen van een Data Warehouse. Als het frame in een andere context wordt gebruikt, kunnen de zuilen naar behoeven worden aangepast. Gemeenschappelijk voor alle toepassingen van het Metadata Frame is de zuil '*Corporate fact base*', die essentieel is voor integratie van alle relevante informatie binnen het bedrijf. Hier verwachten we een bedrijfsbreed gegevensmodel dat vrij is van dubbelzinnigheden en redundanties over alle bedrijfsprocessen heen. In fig. 1 zijn op het conceptuele niveau geen scheidingslijnen tussen de zuilen. Op dit niveau wordt alle relevante informatie immers opgesplitst in elementaire feiten, uitgedrukt in zinnen die geheel vrij zijn van niet-conceptuele informatie of constructie. Deze zijn daarom precies hetzelfde in iedere zuil.

Laten we kijken, hoe alles in zijn werk gaat in de praktijk. We beginnen met fase 1 van de Metadata Frame methode, de voorbereidende fase. Hierin vormen we een beeld van het totale gebruikersdomein, splitsen dit op in gebieden die gescheiden kunnen worden afgehandeld in een evolutionaire aanpak, en voorzien deze van een prioriteit om te bepalen welk gebied we het eerst afhandelen. Het resultaat is een spreadsheet met vijf lijsten:

1. De betrokken organisatie-eenheden.
2. De betrokken processen.
3. De vereiste en gewenste rapporten.
4. Alle relevante performance en key-performance indicatoren (PI's en KPI's).
5. Globaal vastgestelde onderwerpen die in verband kunnen worden gebracht met de (K)PI's, en kandidaat-dimensies zijn in een dimensioneel ontwerp.

De spreadsheet bevat verder een aantal matrices, die aangeven welke organisatie-eenheden, rapporten, (K)PI's, en onderwerpen betrokken zijn bij welke processen, en matrices die het verband aangeven van de (K)PI's met de rapporten en de kandidaat-dimensies. Dit geeft op voorhand een redelijk beeld van de Data Marts die naar voren zullen komen uit het data modelleringsproces. In feite zijn deze matrices het equivalent van een matrix die Ralph Kimball de "Data Warehouse Bus Architecture matrix" noemt [3]: een matrix van processen tegen kandidaat-dimensies. Daarom noemen we deze spreadsheet vaak de '*Bus matrices*'. Fase 1 van de Metadata Frame methode lijkt op een klassieke informatieanalyse, maar het resultaat is geen formeel rapport, maar een helder beeld van alles wat we moeten modelleren in fase 2 van de methode.

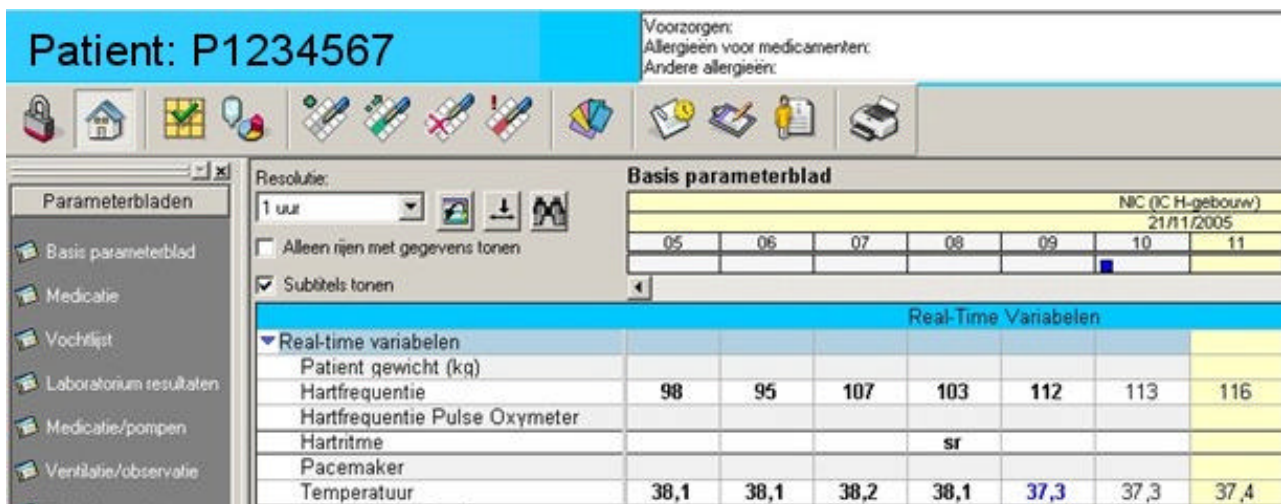
Aan het eind van fase 1 toonde de lijst van rapporten voor het Erasmus MC de vereiste rapporten voor de Zorginspectie, rapporten over patiëntenlogistiek per afdeling en per specialisme, en ad-hoc rapporten voor de medische informatie afdeling, de financiële afdeling, en onderzoek. De lijst van PI's en KPI's was indrukwekkend lang, en bevatte standaard items voor patiëntenlogistiek, zoals aantal ligdagen, opnamen, ontslagen, overplaatsingen van en naar andere afdelingen, overnames van en door andere specialismen, polikliniekbezoeken, geannuleerde bezoeken, en niet-komers. Voor de IC's bevatte de lijst items zoals het aantal positieve bloedkweken, beademingsdagen en uren, IC-opnamen, IC-ontslagen, sterfgevallen, afgezegde operaties, eerste hulp bypass uren, bloedtransfusies, medicaties, en scores zoals de APACHE score<sup>2</sup> en TISS-score<sup>3</sup>. Als kandidaat-dimensies kwamen naar voren: datum, tijd, patiënt, afdeling, specialisme, specialist, behandelend arts, diagnose, type apparaat, medische staf, medicatie, etc. Ten slotte maakten we ook een lijst van attributen die voorkwamen op de vereiste rapporten voor de Zorginspectie.

Fase 2 van onze methode omvat de echte gegevensmodellering met behulp van FCO-IM [4]. Dit vereist het verzamelen en ontwerpen van concrete voorbeelden van alle relevante informatie die in fase 1 is geïdentificeerd. Figuur 2, 3 en 4 tonen voorbeelden daarvan voor de meting van real-time variabelen, laboratoriummetingen, en beademing. De eerste twee komen direct uit het medische bronsysteem. De derde is een voorbeeld dat speciaal voor het modelleren werd ontworpen op basis van analyse van de in fase 1 gevonden PI's en KPI's voor beademingsuren. Ook hier leverden onze medische experts sterke en beslissende bijdragen aan de analyse. Voor patiëntenlogistiek en de IC's samen verzamelden we 15 verschillende sheets met concrete voorbeelden uit de medische bronsystemen en ontwierpen we nog eens 28 sheets om alle relevante informatie in de spreadsheet van fase 1 af te dekken.

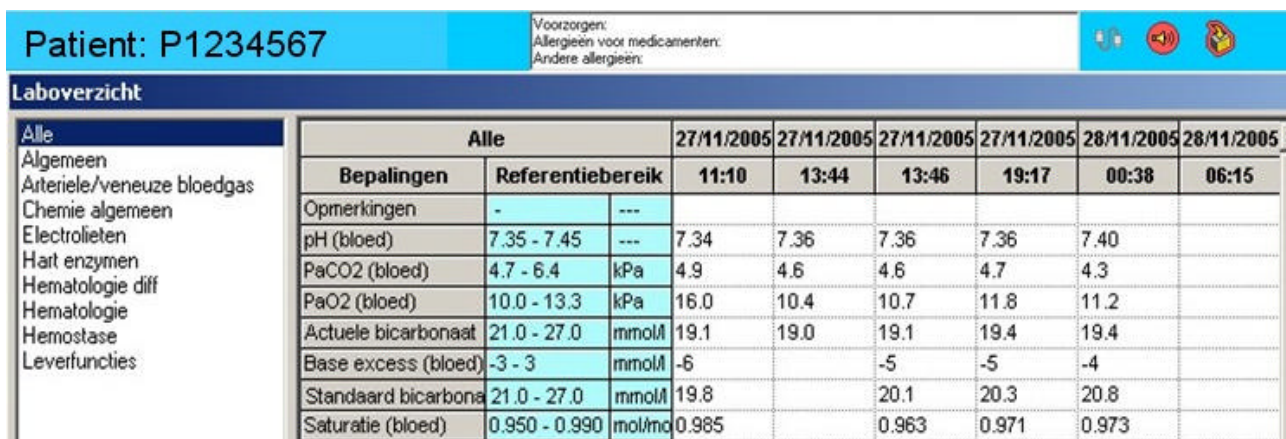
---

<sup>2</sup> Acute Physiology And Chronic Health Evaluation: een maat voor de overlevingskans van patiënten, berekend aan de hand van een verzameling fysiologische variabelen verzameld tijdens de eerste 24 uur op de IC.

<sup>3</sup> Therapeutic Intervention Scoring System: een getal dat de werkdruk van de IC verpleegkundige staf aangeeft.



Figuur 2: Voorbeelden van de meting van real-time variabelen uit het medische bronsysteem.



Figuur 3: Voorbeelden van de resultaten van laboratoriummetingen uit het medische bronsysteem.

IC-datum	Uur (0-23)	Device-type	Beademings- starttijd	Beademings- eindtijd	Patient	Afdeling	Beademings- startdatum	Beademings- starttijd	Aantal uren	Start beademings- periode	Einde beademings- periode
18/11/2005	23:00	BA1	23:35	00:00	1111111	ICP1	18/11/2005	23:35	0.42	1	0
19/11/2005	0:00	BA1	00:00	01:00	1111111	ICP1	18/11/2005	23:35	1	0	0
19/11/2005	1:00	BA1	01:00	02:00	1111111	ICP1	18/11/2005	23:35	1	0	0
19/11/2005	2:00	BA1	02:00	02:22	1111111	ICP1	18/11/2005	23:35	0.37	0	0

Figuur 4: Ontworpen concrete voorbeelden van performance indicatoren voor beademing.

De volgende stap in het proces bestaat uit het verwoorden van alle relevante feittypen die in de concrete voorbeelden te zien zijn. Dit gebeurde samen met als medische experts: de Kinderarts-intensivist en Chief Medical Information Officer, Dr. Jan Hazelzet, een IC-verpleegkundige, Guido Lansbergen, en een promovendus, Marleen de Mul. Voor wat betreft de feiten zichtbaar in fig. 2 waren de experts het met elkaar eens, dat deze correct verwoord waren met zinnen als:

- Er is bij Patiënt P1234567 op 21-11-2005 om 9:00:00 uur voor Hartfrequentie een waarde van 112 slagen / min. vastgelegd.
- Er is bij Patiënt P1234567 op 21-11-2005 om 7:00:00 uur voor Temperatuur een waarde van 38.2 °C vastgelegd.

De feiten zichtbaar in fig. 3 werden volgens de experts correct weergegeven door zinnen als:

- De labbepaling bij Patiënt P1234567 op 28-11-2005 om 00.38 uur voor Standaard bicarbonaat (bloed) had als resultaat 20.8 mmol / l.

- De referentie ondergrens van Patiënt P1234567 op 28-11-2005 om 00.38 uur voor Standaard bicarbonaat (bloed) was 21.0 mmol / l.
- De referentie bovengrens van Patiënt P1234567 op 28-11-2005 om 00.38 uur voor Standaard bicarbonaat (bloed) was 27.0 mmol / l.

Voor de feiten zichtbaar in fig. 4, ten slotte, keurden de experts zinnen als deze goed<sup>4</sup>:

- Het beademingsuurfragment van patiënt 1111111 beginnend op 18-11-2005 om 23:35 uur hoort bij het uur beginnend om 23:00 uur.
- Tijdens het beademingsuurfragment van patiënt 1111111 beginnend op 18-11-2005 om 23:35 uur werd devicetype BA1 gebruikt.
- Tijdens het beademingsuurfragment van patiënt 1111111 beginnend op 18-11-2005 om 23:35 uur werd de patiënt verzorgd op afdeling ICPI.
- Markeert het beademingsuurfragment van patiënt 1111111 beginnend op 18-11-2005 om 23:35 uur de start van een Beademingsperiode? 1.

Zodra de feiten geformuleerd waren, bepaalden we de interne structuur van deze feiten. Ook hierbij konden de medische experts ons uitstekend helpen. Voor de feiten uit fig. 2 waren zij het erover eens, dat het allemaal voorbeelden zijn van een feittype dat bestaat uit een real-time meting en het resultaat daarvan. De real-time meting is de meting van een real-time variabele, geïdentificeerd door zijn naam, op een gegeven moment in het bestaan van een patiënt. Het resultaat bestaat uit een waarde met een geschikte eenheid. Figuur 5 toont de ontledingsboom of “expression tree” voor dit feittype in CaseTalk™ [5], de case-tool die we gebruiken voor de analyse van de feiten. Deze ontleding bepaalt grotendeels de structuur van het op te leveren model. Meer informatie over dit proces is te vinden in [4].



Figuur 5: ‘Expression tree’ in CaseTalk™ voor het feittype “Real-time meting met resultaat”.

Zodra alle relevante feiten in de sheets met concrete voorbeelden verwoord zijn in zinnen goedgekeurd door de medische experts, en op dezelfde manier zijn geanalyseerd als in fig. 5, kan CaseTalk uit al deze feiten een volledig informatiemodel afleiden met behulp van een exact algoritme. Dit algoritme wordt gedicteerd door ‘predicatenlogica’, de wiskundige theorie achter FCO-IM. Het produceert een relationeel model dat aan alle normalisatieregels van Codd voldoet met het minimum aantal tabellen dat nodig is om alle informatie te bevatten die correspondeert met de relevante feiten: niets meer en niets minder. Het feittype in fig. 5 verschijnt in dit model als onderdeel van de tabel getoond in fig. 6. In deze figuur komt het laatste veld in de tabel Real\_Time\_meting voort uit een feittype waarvan de eerste zin blijkbaar was: “Is het resultaat voor Hartfrequentie bij Patiënt P1234567 op 21-11-2005 om 9:00:00 uur gevalideerd? 1.” Deze zin verwoordt een feit getoond op een niet zichtbaar deel van de voorbeeldsheet in fig. 2, waar was aangegeven dat de dienstdoende verpleegkundige staf deze gegevens gevalideerd had.

Zoals blijkt uit de foreign-key toewijzingen in fig. 6, produceert het algoritme ook de tabellen Patiënt, Dag, en Tijdstip die voorkomen in de expressieboom in fig. 5. Deze tabellen hebben als primary-key Patientnummer, Datum, en Tijd, en bevatten vele andere velden voortvloeiend uit relevante feiten in alle verzamelde voorbeelddocumenten. Niet verrassend maar beslist plezierig is het feit, dat de tabellen Patiënt, Dag, en Tijdstip ook voorkwamen als kandidaat-dimensies in de spreadsheet van fase 1 van onze methode. Daar werd de tabel Tijdstip simpelweg aangeduid met Tijd. In dezelfde geest leiden analyses als in fig. 5 van de feittypen die horen bij de figuren 3 en 4 tot de tabellen “Laboratorium\_meting” en “Beademingsuurfragment”.

<sup>4</sup> De volledige uitwerking van de voorbeelddocumenten en hun verwoordingen is te vinden op [www.array.nl](http://www.array.nl).

**Tabel: Real\_Time\_meting**

RT_variabelenaam (RT_variabelenaam)	Patientnummer (Patientnummer)	Datum (Datum)	Tijd (Tijd)	Waarde (Waarde)	Eenheid (Eenheid)	Gevalideerd (Boolean)
NN	NN	NN	NN	NN	NN	NN
PK	PK	PK	PK			
Hartfrequentie	P1234567	21-11-2005	9:00:00	112	slagen / min.	1
Temperatuur	P1234567	21-11-2005	7:00:00	38.2	°C	1

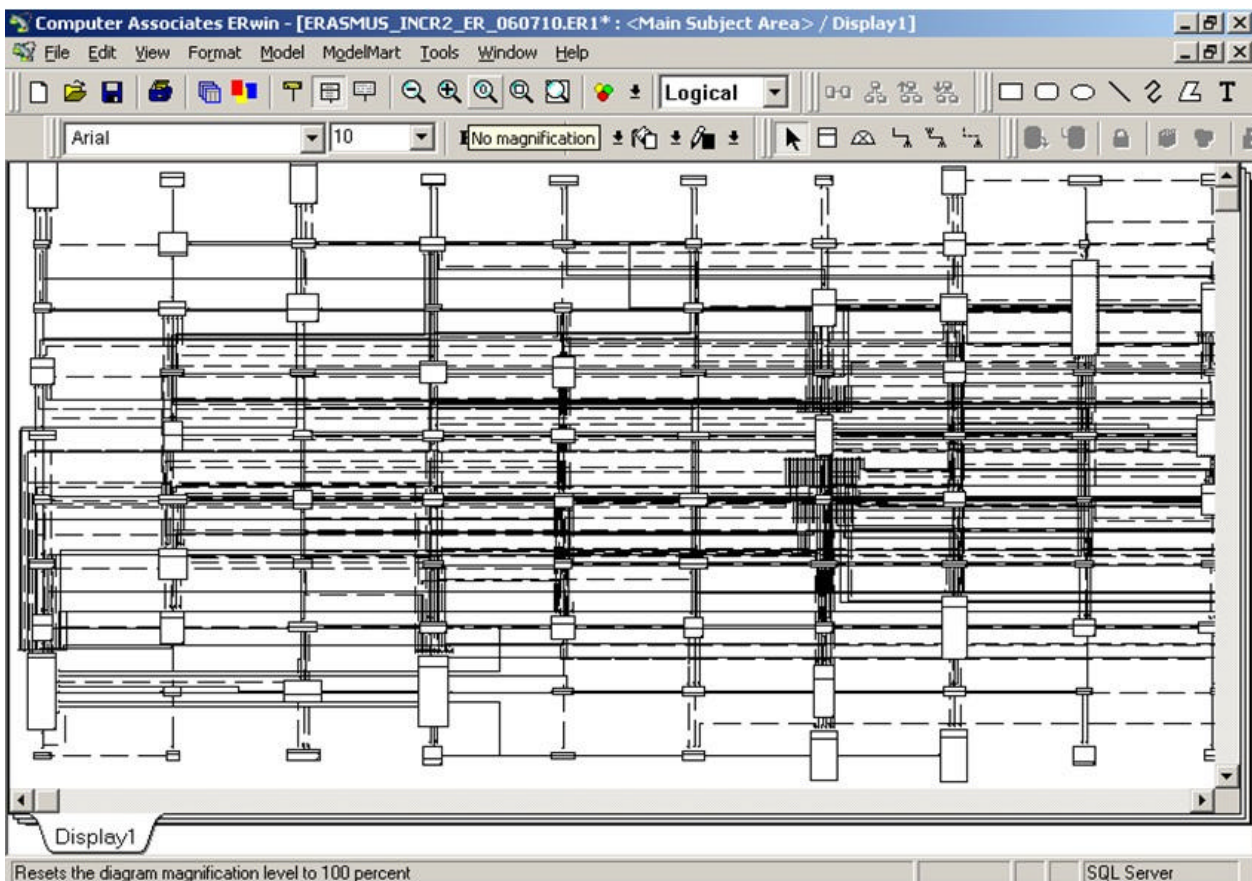
**FEITTYPE EXPRESSIONS:**  
F1: "Er is bij Patiënt <Patientnummer> op <Datum> om <Tijd> uur voor <RT\_variabelenaam> een waarde van <Waarde> <Eenheid> vastgelegd."  
F2: "Is het resultaat voor <RT\_variabelenaam> bij Patiënt <Patientnummer> op <Datum> om <Tijd> uur gevalideerd? <Gevalideerd>."

**FOREIGN KEYS:**  
Real\_Time\_meting (RT\_variabelenaam) → RT\_variabele (RT\_variabelenaam)  
Real\_Time\_meting (Patientnummer) → Patiënt (Patientnummer)  
Real\_Time\_meting (Datum) → Dag (Datum)  
Real\_Time\_meting (Tijd) → Tijdstip (Tijd)

*Figuur 6: De tabel Real\_Time\_meting voortvloeiend uit het model creërend algoritme van CaseTalk.*

Als CaseTalk zijn algoritme toepast, wordt een "logisch data model" in optimale normaalvorm geproduceerd voor de Corporate Fact Base, maar dit zit nog steeds in de repository van CaseTalk. Een ander hulpmiddel in de volledige toolset van het Metadata Frame, de ERM-Bridge, kan het model in zijn geheel vanuit CaseTalk naar diverse ER/UML-omgevingen transporteren. Dit hulpmiddel is oorspronkelijk ontwikkeld op de "Hogeschool van Arnhem en Nijmegen", die nog steeds de bakermat van FCO-IM is en onze partner bij het uitdenken van vele van de ideeën die in ons Metadata Frame worden toegepast.

Na het transporteren van het model, is dit beschikbaar in de gekozen ER-tool zonder verlies van enige belangrijke informatie over het model. In het bijzonder is alle semantiek die aan het model ten grondslag ligt, overgebracht in de ER-repository in de vorm van commentaren bij de attributen. De precieze betekenis van elk veld in het model kan daarom nog steeds begrepen worden in termen van de oorspronkelijke zinnen.



*Figuur 7: Het volledige ER-model voor het Erasmus MC zoals het door de algoritmen is gegenereerd.*

Het volledige model in de ER-tool, dat door de algoritmen is gegenereerd, is zichtbaar in fig. 7. En inderdaad, het is een heel groot model. Dit is precies waarom het niet met de hand moet worden gemaakt in IT-gedreven top-down benadering, waarbij de experts niet serieus betrokken zijn. Zelfs in het onwaarschijnlijke geval dat een dergelijke benadering tot een volledig correct model van de communicatie leidt, zijn mensen niet in staat het model volledig correct in te voeren. Een dergelijk model kan beter gegenereerd worden door een strak algoritme vanuit een conceptueel model dat volledig gevalideerd is door experts, zoals in onze aanpak het geval is. Het algoritme is niet in staat om fouten te maken. Daarom weten we domweg, dat het resulterende ER-model goed is. Of zelfs beter: dat het *perfect* is.

## Ontrafeling van de Babylonische spraakverwarring

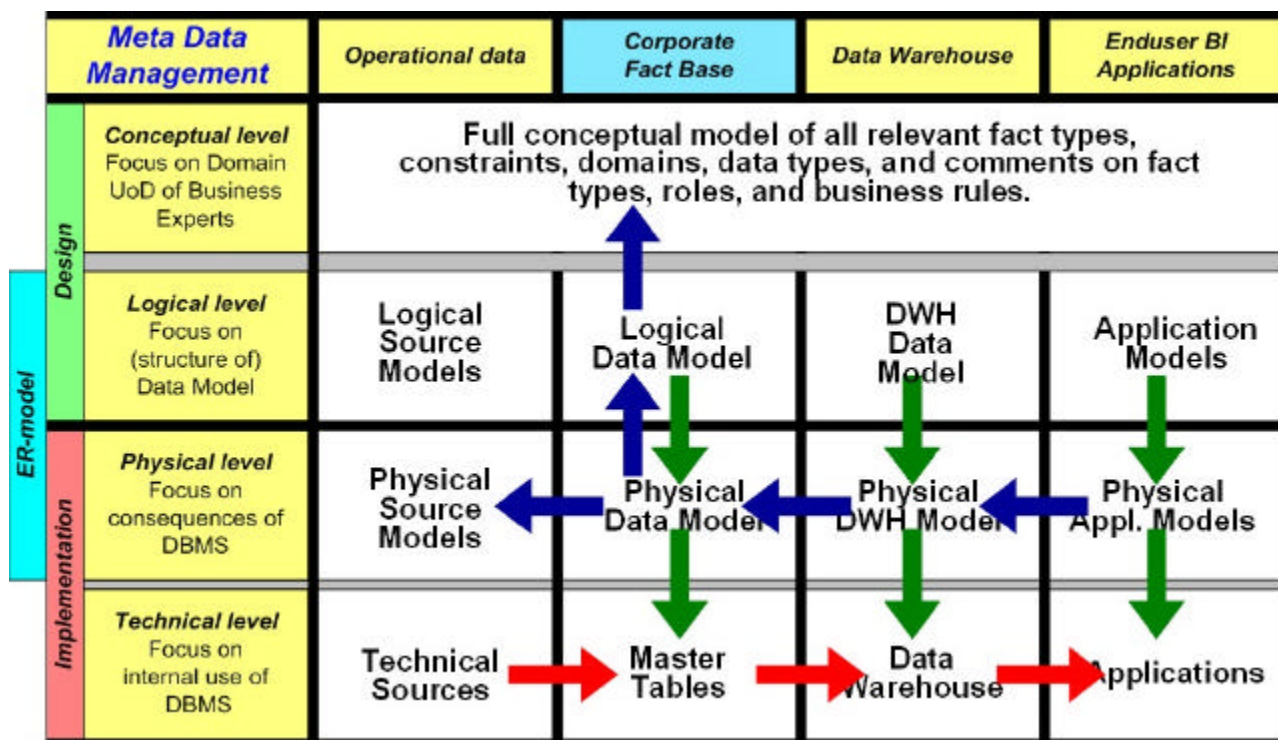
In principe kan onze modelleringsaanpak in fig. 1 ook worden gezien als een ‘Plan-Do-Check-Act’ cyclus. Zodra deze cyclus is afgerond, kunnen we de modellen creëren die in de Data Warehouse omgeving geïmplementeerd worden. De feitelijke validatie van deze modellen is dan al achter de rug. Dit betekent, dat onze Metadata Frame methode voldoet aan de volgende eisen:

- *Doelgerichtheid:*
  - We modelleren niet de werkelijkheid, maar de *communicatie* van de business experts over hun werkelijkheid. Dit resulteert in feitexpressies in de taal van de business experts zelf.
  - We modelleren alle conceptuele aspecten van de communicatie *zonder vooroordelen of beperkingen*.
- *Doeltreffendheid:*
  - De business experts zijn in staat te valideren dat hun communicatie volledig en correct gemodelleerd is, zonder zelf naar enig gegenereerd data model te kijken *en zonder hun kostbare tijd te verspillen aan voor hen irrelevante en onbegrijpelijke technologische aspecten*.
- *Doelmatigheid:*
  - We slaan alle modelinformatie op in één generieke repository voor alle niveaus, en gebruiken repository transformaties om alle metadata via algoritmen te transporteren naar iedere plek waar deze nodig is voor verdere naslag.

Als IT-ers hebben wij in dit proces de rol van *leverancier*, en met onze methode tonen we volledig respect voor de business experts als onze *klanten*. Doordat we in het ontwerpproces volledige validatie verkrijgen, creëren we *volledige klanttevredenheid*. Dit leidt tot minimale verspilling van tijd en hulpmiddelen in de implementatiefase. Dat blijkt uit de volgende uitspraken gedaan door onze collega’s, die het Data Warehouse bij het Erasmus MC op basis van ons ontwerp gebouwd hebben:

- “Het model is volledig begrijpelijk voor alle afnemers, eindgebruikers en ontwikkelaars gelijk: de zinnen en voorbeelden maken het eenvoudig te communiceren over de te leveren functionaliteit.”
- “De ontwikkelaars worden door de voorbeelden sterk ondersteund bij de bouw: de voorbeelden prikkelen hun voorstellingsvermogen op de juiste manier.”
- “De gehele toolset van het Metadata Frame werkt als communicatiemiddel tussen de medische experts en de IT-staf.”
- “De metadata zijn moeiteloos beschikbaar te maken bij de attributen zowel in de ETL als in de front-end applicaties.”
- “De gemodelleerde stermodellen bevatten *alle* rapportagebehoeften: het bouwen van de Business Objects universes is kinderspel.”

Voor wie het Metadata Frame kent, is dit geen verrassing: *alle* rapportagebehoeften dienden als invoer voor een proces dat volledige validatie garandeert. Verder zijn de functionele stappen in het implementatieproces (zie de pijlen in fig. 8) maximaal geautomatiseerd, waarmee de kans op fouten is geminimaliseerd. De blauwe pijlen geven de processen weer, die door onze metadata tool MetaMorfose worden uitgevoerd. Deze tool legt automatisch alle informatie over de modellen vast, die voor de diverse zuilen zijn ontworpen, en koppelt deze modellen grotendeels geautomatiseerd aan elkaar. Dit ondersteunt krachtig de rode pijlen, die de ETL-processen weergeven, en evenzo de geautomatiseerde levering van relevante metadata aan eindgebruikers in hun front-end applicaties. Dit geldt in het bijzonder voor alle semantiek die aan de velden in deze applicaties ten grondslag ligt. Zoals we lieten zien in [6], hebben de zinnen die deze velden genereerden hebben veel meer betekenis voor een eindgebruiker dan standaarddefinities van deze velden. Definities neigen de aandacht te trekken naar de *geldigheid* van de inhoud van velden, terwijl de concrete semantiek de aandacht trekt naar de *juistheid* van deze inhoud. En gezien vanuit informatiekwaliteit is het laatste veel belangrijker dan het eerste.



Figuur 8: Het Metadata Frame en zijn functioneel gebruik in de implementatieprocessen.

Toen de eerste resultaten van de applicatieontwikkeling beschikbaar kwamen, verzekerden de betrokken medische experts ons al spoedig, dat zij met deze applicaties alle wensen aan het begin van dit artikel konden vervullen, en dat zij er inderdaad erg verguld mee waren. Dit werd nog verder opgeluisterd door de gelukkige uitkomst die hieronder is getoond: Het Erasmus MC en Atos Origin hebben met dit project de Computable Award 2007 gewonnen voor het beste IT-project in de Zorg.

### Referenties

1. Database Magazine, Jaargang 11, no. 8 (december 2000), P.W.F. Alons.
2. Database Magazine, Jaargang 12, no. 1 (januari 2001), P.W.F. Alons.
3. The Data Warehouse Lifecycle Toolkit, Ralph Kimball, Laura Reeves, Margy Ross, Warren Thornthwaite, (Wiley Computer publishing, 2002), ch. 7.
4. Volledig Communicatiegeoriënteerde Informatiemodellering, G. Bakema, J.P. Zwart, H. van der Lek, Kluwer BedrijfsInformatie, 2002
5. CaseTalk, Bommeljé, Cromptvoets, en Partners, <http://www.casetalk.com/php/>.
6. Database Magazine, Jaargang 17, nr. 2 (april 2006), P.W.F. Alons en R.G. Arntz.

### Over de Auteur

Dr. Peter W.F. Alons is senior consultant bij Atos Origin/BI-CRM en al ruim vijftien jaar betrokken geweest bij een groot aantal Business Intelligence en Data Warehouse projecten bij diverse bedrijven in diverse branches.

### Contact

Peter Alons: [peter.alons@atosorigin.com](mailto:peter.alons@atosorigin.com)

Atos Origin: [www.atosorigin.com](http://www.atosorigin.com)



**De Computable Award 2007 for de best IT-project in Healthcare: Erasmus MC en Atos Origin.**