

Productfamilie wil een totale metadata-oplossing bieden

# Ontwikkelingen rond ERwin

Theo van Westrienen

**Het is de laatste jaren nogal stil op het gebied van datamodelleringstools. Er viel de afgelopen jaren weinig anders te lezen dan incidentele aankondigingen van nieuwe versies. Dat betekent niet dat er geen tools meer gebruikt worden, of dat de ontwikkelingen stil staan. Integendeel, de markt is volop in beweging: met enige regelmaat treden nieuwe spelers toe.**

Veelal betreft dat kleinere spelers, met tools die zich beperken tot fysiek databaseontwerp en eenvoudige reverse engineering. Ook is er een herintreder: IBM stort zich – na jarenlange afwezigheid – weer op deze markt, en zet sinds ruim twee jaar vol in met Data Architect. Daarmee is IBM volgens Gartner een uitdager voor de grote drie: ER/Studio, PowerDesigner en marktleider ERwin. Dit artikel gaat in op de stand van zaken rond ERwin, met een korte terugblik, gevolgd door een overzicht van recente ontwikkelingen en concrete toekomstplannen.

## Terugblik

Het is 1989 wanneer ERwin, ontwikkeld door LogicWorks, op de markt komt. In 1998 neemt Platinum ERwin over, waarna nauwelijks een jaar later Computer Associates, het huidige CA, op haar beurt Platinum overneemt. Deze overnames vormen in eerste instantie een rem op de ontwikkeling: voor de eerste nieuwe release die CA in 2000 op de markt brengt, is de basis nog gelegd door LogicWorks. De organisatorische en personele wisselingen hebben tot gevolg dat er onvoldoende grip is op richting en stabiliteit van het product. Het duurt feitelijk tot 2004 voor release 4 met versie 4.1.4 een zodanig stabiel niveau bereikt dat bestaande gebruikers massaal overstappen.

Dan zijn de voorbereidingen voor de volgende major release in volle gang. CA realiseert zich dat het anders moet: de ervaring met release 4 leert dat de code base niet meer rendabel onderhoudbaar is en vernieuwingen belemmert. Functies zijn veelal op meerdere plaatsen in de code aanwezig, hetgeen foutgevoelig is en zorgt voor inconsistentie. Een crime om bestaande functionaliteit aan te passen, dan wel nieuwe toe te voegen.

## Ontwikkelingen

Om door te kunnen ontwikkelen moet de interne architectuur op de schop. CA maakt heel even een pas op de plaats en begint

van voren af aan met de opbouw van een volledig nieuwe code base. Centraal hierin staat de zogenaamde Shared Modeling Engine. Niet alleen ERwin maakt hier gebruik van, ook de andere producten uit de suite zoals ModelManager doen dat. Daarnaast kiest CA ervoor om minder zelf te bouwen door samenwerkingsverbanden aan te gaan. Het eerst zichtbare resultaat daarvan vormt de samenwerking met Meta Integration, waarover later meer.

De eerste versie van de nieuwe release, R7, is in 2006 beschikbaar, en daarmee het eerste deel van de vernieuwde code. Besloten is om de vernieuwingen gefaseerd door te voeren, verspreid over de verschillende subreleases van R7. Doel is om de fase van volledige code base vervanging met R8, gepland voor de eerste helft van 2010, af te sluiten.

Deze gefaseerde oplevering komt doordat naast de technologische aanpassing ook op functioneel gebied veel vernieuwingen plaats vinden. Hieronder volgt een overzicht van de belangrijkste zaken die vanaf 2006 tot en met januari 2009 met het verschijnen van 7.3 zijn toegevoegd.

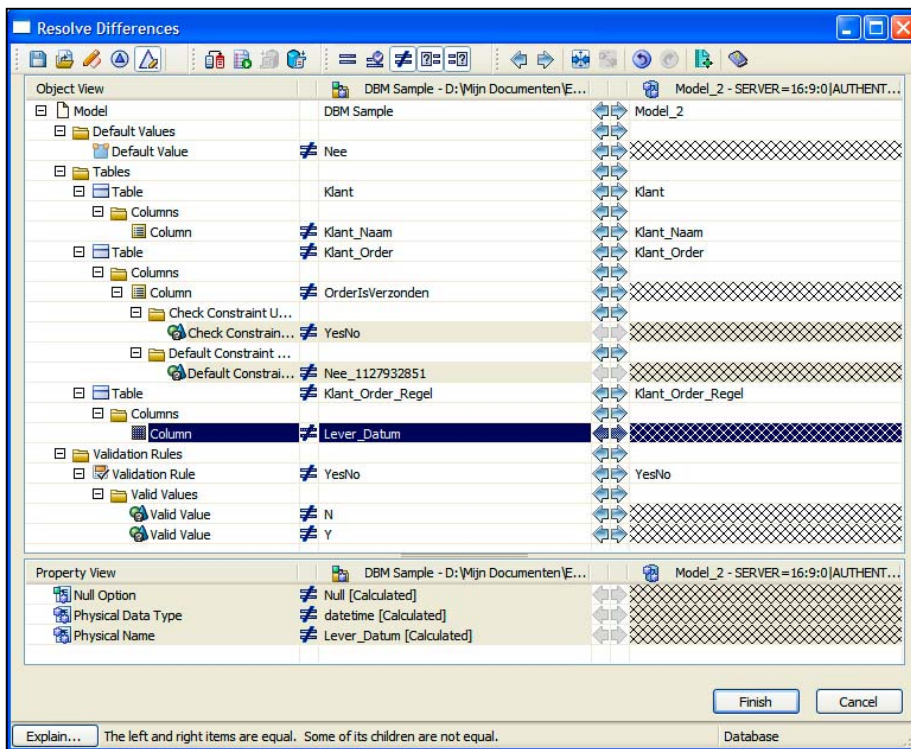
### Complete Compare.

De Complete Compare (CC) functie vormt de grootste kracht van ERwin. Met deze functie kun je modellen onderling vergelijken. Een model kan hier een logisch dan wel een fysiek model zijn, maar ook een database. CC bestond al, maar is sterk uitgebreid. CC maakt geen onderscheid meer tussen de verschillende typen vergelijkingen: er is één interface die voor elke vorm van vergelijking van toepassing is.

Wanneer ModelManager aanwezig is dan zijn de vorige versies te vergelijken. En wanneer je met meerdere gebruikers tegelijkertijd in hetzelfde model werkt, dan kun je met dezelfde functie de eventuele conflicterende wijzigingen vergelijken en verwerken.

Een optionele wizard loodst je stap voor stap door het CC proces. Eerst selecteer je de twee te vergelijken modellen. Als een of beide daarvan een database is, dan vindt vervolgens een impliciete Reverse Engineering plaats: via een native of ODBC connectie met de betreffende database lees je de geselecteerde eigenschappen van de database. ERwin zet deze intern om in een model.

In de volgende stap selecteer je de objecteigenschappen waarop je gaat vergelijken. De keuze is hierbij zeer groot. Voor vrijwel



**Afbeelding 1:** Resolve Differences scherm.

elke modeleigenschap is te kiezen of ERwin wel of niet op dat aspect vergelijkt. Alleen verschillen tussen geselecteerde eigenschappen van de twee modellen komen in beeld voor verdere verwerking. Zo kun je bijvoorbeeld bepalen dat je voor kolommen wel het datatype, maar niet de defaultwaarde vergelijkt. Dan volgt de selectie van de objecten in de twee modellen: welke objecten moeten deel uitmaken van de vergelijking? Men kan hier elke subset van entiteiten, attributen, relaties, schema's enzovoort samenstellen. Je kunt kiezen om in een keer alle objecten van een of meer specifieke Subject Areas te selecteren. In praktijk is dat wat je vaak doet: na werkzaamheden in de huidige Subject Area, oftewel de huidige subset, wil je deze vergelijken met een ander model of met de database. De verschillende geselecteerde opties, zowel voor Reverse Engineering als voor de te vergelijken eigenschappen, zijn in de vorm van sets op te slaan voor hergebruik. Datzelfde geldt voor de complete sessie, dus inclusief optie sets, modelnamen en locatie. In een laatste scherm zijn er nog wat specifieke opties en dan voert ERwin de eigenlijke vergelijking uit. Dat leidt tot het Resolve Differences scherm, zie afbeelding 1, het scherm waarop ERwin de verschillen tussen de modellen toont en je bepalen kunt per verschil of groep van verschillen wat je wilt doen: promoten van het linker- naar het rechtermodel of vice versa, dan wel negeren. Uiteindelijk genereert ERwin desgewenst een Alter Script waarmee de wijzigingen in de eventuele database zijn uit te voeren. Modelwijzigingen brengt ERwin direct aan.

#### DBMS Support.

Ondersteuning van de diverse DBMS'en is verbeterd door de

introductie van externe, aanpasbare database support templates. Aanpassing vindt plaats in de nieuwe Integrated Development Environment, welke gebaseerd is op de scripttaal TLX. Hiermee kan men alle modeleigenschappen benaderen en gebruiken. Het is de bedoeling dat deze IDE in volgende releases uitgroeit tot volledige vervanger van de bestaande macro- en scriptomgeving van ERwin.

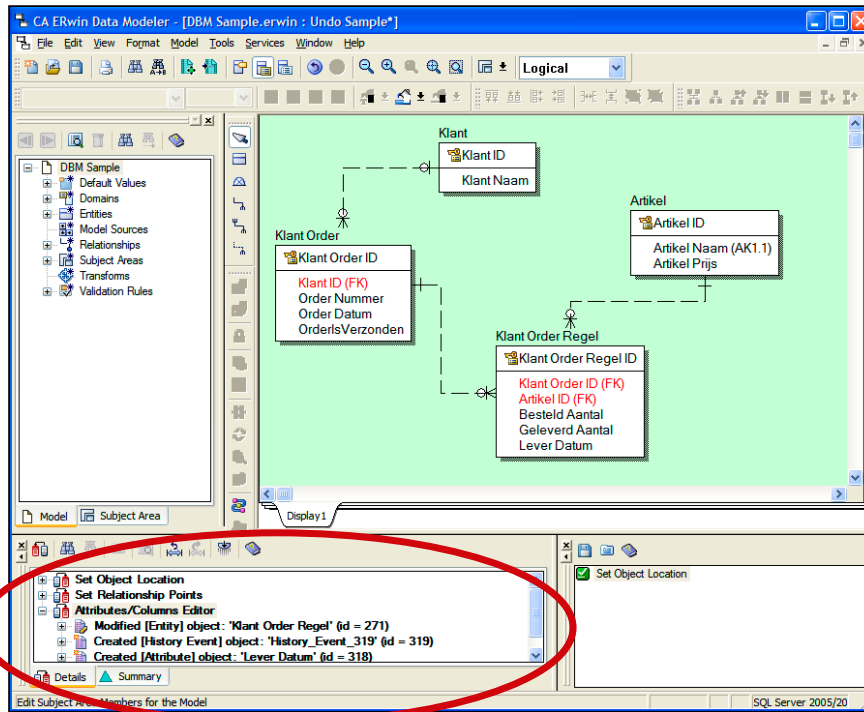
#### Integratie.

De samenwerking met Meta Integration betekent dat CA standaard de import- en exportbruggen van Meta Integratie meeleverd bij ERwin. Dat betekent dat je kunt importeren van en naar een scala van relevante producten en bestandsformaten.

Voorbeelden hiervan zijn: IBM InfoSphere producten zoals Data Architect en Data Stage, Cognos, Rational Rose en Telelogic. ER/Studio, Powercenter, MS Office, OMG (CWM en UML), Business Objects en Oracle Designer en nog veel meer.

#### ODBC Rapportage.

Voor rapportage kende ERwin tot voor kort twee beperkte hulpmiddelen: de Data Browser en de Report Template Builder. Beide zijn volledige zelfgebouwde rapportgeneratoren, met als belangrijkste beperking dat slechts een subset van alle modeleigenschappen beschikbaar is voor rapportage. De ontwikkeling van beide is stopgezet, en in plaats daarvan is nu het ERwin model zelf benaderbaar via een native ODBC connectie, die het volledige interne datamodel ontsluit. Om daar optimaal van te profiteren levert CA ERwin nu standaard met een volledige Crystal Reports licentie.



Afbeelding 2: Action log.

## Zoek en vervang.

De nieuwe zoek- en vervangfunctie doorzoekt alle model-metadata op basis van de opgegeven zoekstring, en kan de gevonden strings vervolgens aanpassen.

## Name Hardening.

ERwin kent al sinds jaar en dag de mogelijkheid om naamgevingstandaarden toe te passen, maar pas nu is de zogenaamde hardening toepasbaar. Daarmee voorkom je voor (groepen van) objecten dat de naam wijzigt wanneer de naamgevingstandaard die voor dat object van toepassing is gewijzigd.

## Undo/Redo.

Het heeft lang geduurd voordat de Undo optie beschikbaar kwam, maar nu is de functionaliteit wel volledig: elke individuele actie gedurende een gehele sessie is ongedaan te maken – en desgewenst weer opnieuw uitvoerbaar. Dit is onafhankelijk van andere acties of van volgorde, en van toepassing op elk type wijziging. Het blijft ook mogelijk wanneer het model tussentijds wordt opgeslagen. In afbeelding 2 is het Action Log, het hart van deze functionaliteit, omcirkeld.

## ERwin Suite

Met R7 verandert niet alleen ERwin zelf maar ook de suite rond ERwin. De positionering van de ERwin familie wijzigt van database ontwerptool naar totale ontwerp oplossing voor het gebied van Data Management, zie afbeelding 3. Dat betekent dat de ERwin productfamilie gestaag uitbreidt: naast bekende componenten als Model Manager en Process Modeler zijn de afgelopen jaren al de Saphir option en Model Validator toegevoegd. Model Manager omvat het terrein van versiebeheer en multi-

user modelgebruik, en met Process Modeler ontwerp je Business Processen. De Saphir Option ondersteunt metadata-extractie van ERP-systemen zoals SAP, PeopleSoft en Siebel, onder andere ten behoeve van systeemintegratie. Model Validator tenslotte controleert en valideert de kwaliteit van logische en fysieke datamodellen, geeft advies voor verbeteringen en kan die ook implementeren in het model.

De nieuwste toevoeging is sinds januari 2009 ERwin Data Profiler. Met Profiler analyseer je bestaande gegevensbronnen en bepaal je inhoud, structuur en kwaliteit van de gegevens. Profiler komt voort uit een samenwerking met Exeros.

## Roadmap

De ambitie van CA is af te leiden uit de roadmap: met de ERwin productfamilie een totale metadata-oplossing bieden. Naast brede RDBMS-ondersteuning impliceert dat automatiseren van de werkzaamheden op het terrein van data administratie, -analyse en -modellering.

De belangrijkste verbeteringen voor de eerstvolgende release, R8, aangekondigd voor de 1e helft van 2010 zijn:

*Database Support.* Volledige support DB2/UDB (z/OS en LUW), Teradata 13 en SQL Server 2008. In het bijzonder de support van DB2 is hier relevant – de huidige versie ondersteunt v8 weliswaar volledig, maar nog niet alle nieuwe functionaliteit van v9; *Verbeterde Lay-out.* De lay-out ondergaat een metamorfose, en gaat nu onder andere stylesheets, verbeterde modelnavigatie en flexibele zoom-mogelijkheden bieden. De verwachting is dat CA dit realiseert door middel van een nieuw samenwerkingsverband met een third party; *Bulk editor.* Geïntegreerde functionaliteit – vergelijkbaar met een spreadsheet – waarmee je zelf een configureerbare view op de

modelmetadata kunt opzetten, waarin je die gegevens desge-  
wenst kunt muteren.

Overige verwachte verbeteringen en uitbreidingen voor deze  
release betreffen onder andere spellingscontrole, wizards, volle-  
dige integratie van grafische modelpresentaties met Crystal  
Reports, transformaties en historie van events.

Voor de daarop volgende releases richten de verbeteringen zich op:

- Volledige RDBMS-support voor huidige en voorgaande versies  
van alle ondersteunde databases;
- Verbeterde ondersteuning van industriestandaarden waaron-  
der MS Windows;
- Vergaande configureerbaarheid van defaults, voorkeursinstel-  
lingen en dergelijke;
- Nesting van modelbibliotheken en subject areas;
- Productiviteit en performance;
- Delen van objecten tussen meerdere modellen, inclusief cross  
reference faciliteiten;
- Uitbreiding van ondersteuning van views, dimensioneel  
modelleren;
- Command line interface en overige uitbreidingen van API  
extensies.

## Sterke en zwakke kanten

Op basis van praktijkervaring laat ERwin zich kenmerken door  
een aantal sterke en zwakke aspecten. Het eerste zwakke aspect  
is hierboven al genoemd: de ondersteuning van DB2 loopt achter.  
CA heeft weliswaar op dit terrein een stap gezet door DBMS-  
ondersteuning te externaliseren op basis van templates, zodat  
je zelf aanpassingen kunt doorvoeren indien nodig, maar dat is  
niet altijd voldoende. Voor diverse DBMS'en is de ondersteuning

up-to-date, maar in sommige gevallen, zoals DB2 v9 blijft de  
ondersteuning te lang op het niveau van tolerantie in plaats van  
volledige ondersteuning.

De ondersteuning voor views is een andere zwakke plek. Deze is  
de laatste tien jaar niet gewijzigd, laat staan verbeterd.

Weliswaar kun je elke syntactisch toegestane view in ERwin  
ontwerpen en vastleggen, maar alleen eenvoudige views zijn  
volledig te realiseren met klikken en slepen. Enige complexiteit  
betekent al snel handmatig uitschrijven.

Het Complete Compare proces is een voorbeeld van zowel een  
sterk als zwak aspect: het is een sterk aspect vanwege de zeer  
uitgebreide functionaliteit ervan. Alleen is deze zo sterk uitge-  
breid, dat dit soms ten koste is gegaan van productiviteit en  
performance. Juist omdat je op elke eigenschap van elk object  
binnen een model of database kunt vergelijken, zijn er meer  
handelingen (muis- en toetsenbord klikken) nodig, en meer tijd,  
om een Complete Compare proces te doorlopen.

Het lijkt erop dat CA zich met de introductie van release 7  
destijds meer gericht heeft op modelleren van de uitzondering  
dan op de alledaagse praktijk. Voor veel gebruikers is dat nog  
steeds het vergelijken van een subset van een datamodel met  
een subset van een database, namelijk het deel waar jouw  
project zich op dat moment mee bezig houdt.

Gelukkig zijn er op dit gebied sindsdien al veel verbeteringen  
doorgevoerd, en dat zet zich door in de komende Service Packs  
en releases.

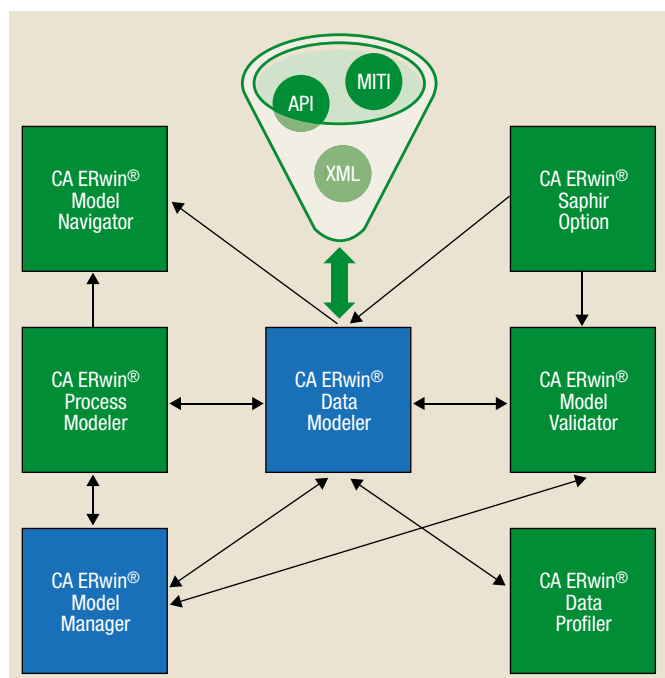
ERwin kent een strak beleid ten aanzien van patch releases en  
Service Packs: elke een tot twee maanden is er een patch  
release, en twee keer per jaar een Service Pack. Dit jaar bijvoor-  
beeld zijn er na de verschijning van release 7.3 eind januari vier  
patch releases geweest en een Service Pack.

Een ander sterk punt is de mogelijkheid om de informatie die in  
een ERwin model zit op allerlei manieren te hergebruiken.  
Natuurlijk kan men standaard met behulp van gegenereerde  
DDL databases genereren en wijzigen. Maar je kunt daarnaast  
met behulp van scripts DML, procedurele code, XML/XSD struc-  
turen en meer genereren. In praktijk betekent dat je bijvoorbeeld  
de volledige data laag voor een applicatie kunt genereren en  
onderhouden.

Tenslotte is er een actieve User Community: wereldwijd zijn er  
meer dan 50 gebruikersgroepen actief, waaronder sinds 2004 ook  
een in de Benelux. Naast het nog vrije jonge ERwin forum op de  
CA support site, is er al 20 jaar een sterke, actieve en onafhanke-  
lijke ERwin nieuwsgroep, bereikbaar via [www.infoadvisors.com](http://www.infoadvisors.com).  
Op deze site zijn diverse nieuwsgroepen te vinden voor data-  
modelleringstools, zoals ER/Studio, PowerDesigner en InfoSphere  
Data Architect, en daarnaast een algemene, toolonafhankelijke  
datamodelleringsgroep.

### Theo van Westrienen

Th.J. van Westrienen is Data Manager bij AEGON Bank N.V. en is  
Voorzitter BeNeLux Modeling User Group.



Afbeelding 3: Overzicht ERwin Suite.