

MetaDimensions wil helpen bij integrale aanpak

KNOWtory, jonge specialist in metadatabeheer

Sandra Gerrits

Sinds kort begint er hoop aan de horizon te gloren voor menige metadata-adept. Langzaam maar zeker ontstaat een geheel nieuwe generatie metadatatools en -applicaties gericht op het management van metadata in zijn geheel: zowel voor systeemgebruikers als technische en functionele gebruikers, en speciaal gericht op alle facetten van metadatabeheer. Een van de nieuwe spelers op de metadatamarkt is MetaDimensions met het product KNOWtory.

Systemen als KNOWtory richten zich op een gat dat in de markt is ontstaan. Traditionele repository's sluiten niet aan op de wensen en eisen van de hedendaagse, mondige en veeleisende gebruiker. Met deze tools wordt beoogd te bereiken wat menige gebruikersorganisatie de afgelopen jaren zelf heeft geprobeerd te bouwen, omdat het niet te koop was: een goede, dynamische, flexibele en vooral toegankelijke integratie van de gewenste en gebruikte metadata. De nieuwe software staat nog aan het begin van haar ontwikkelperiode. Toch kan zij over het algemeen goed met de complexe metadataproblematiek uit de voeten.

KNOWTORY BASICS

Een van de belangrijkste kenmerken van KNOWtory is de mogelijkheid data in elke vorm in te lezen, ongeacht de bron: een MS Excel-spreadsheet, de repository van een dmbs, gegevens uit een CASE-tool, het Web of nog anders. Nu geldt dat voor veel metadata-

Juist in een ingezakte markt is de noodzaak van goed metadatabeheer des te dringender

tatools, maar het grote verschil zit 'm erin dat de structuur van de aangeleverde data in KNOWtory behouden blijft. De gegevens hoeven niet in een vooraf gedefinieerd formaat te worden aangeleverd om te kunnen worden ingelezen. En in de applicatie zelf

wordt de originele opslagstructuur gehandhaafd. Door de gegevens zo fijnmazig mogelijk op te slaan en de inhoud los te koppelen van de semantiek -waarbij gedrag en structuur apart worden vastgelegd- kunnen de gegevens op allerlei manieren gebruikt en weergegeven worden. Hierdoor is het bijvoorbeeld mogelijk de gegevens van een databaseschema te koppelen aan de metadata van de CASE-tool en dat geheel aan te vullen met businessdata uit

Metadata, stiefkindje van de IT

Juist in een ingezakte markt, waardoor veel bedrijven zich gedwongen zien om met minder mensen dezelfde en zelfs verbeterde dienstverlening te verlenen, is de noodzaak van goed metadatabeheer des te dringender. Helaas is 'metadata' nog steeds een stiefkindje van de IT, dat tot voor kort nauwelijks mogelijkheden heeft gehad om op te groeien. Met als gevolg dat menig bedrijf voorlopig nog inefficiënt en onvolledig gebruik maakt van de in de organisatie beschikbare gegevens. Door gebruik te maken van goed georganiseerde metadata, kan een organisatie de enorme hoeveelheid gegevens die zij bijhoudt, inzetten voor optimalisatie van klantrelaties, leverancierrelaties, marketingacties, bedrijfsprocessen en -rapportages en overige informatiegebieden. Metadatabeheer maakt een snellere en makkelijker ontwikkeling van nieuwe applicaties mogelijk. Ook het beheer van applicaties en systemen wordt erdoor vereenvoudigd. Gebruikers krijgen betere, minder omslachtige mogelijkheden om zelf op zoek te gaan naar de informatie die zij nodig hebben, om vervolgens beslissingen te nemen op basis van betere informatie. Maar de uitdagingen op dit gebied zijn nog groot. Dat komt door de complexiteit van metadata en doordat tot op heden slechts kleine successen zijn geboekt om het metadata-universum toegankelijk te krijgen voor de technische-, functionele- en systeemgebruiker. De aloude stelregel 'groot denken, maar met kleine stapjes beginnen' geldt ook voor het opzetten van een metabase en de metadata-architectuur.

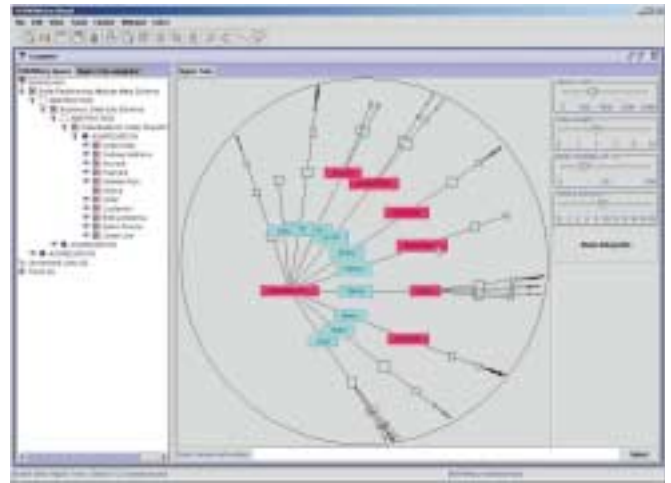
een Excel-bestand. Gegevens van verschillende metadatabronnen zijn gemakkelijk vergelijken, waardoor snel winst valt behalen in ontwikkel- en verbetertrajecten.

Classificatie en het 'instant' maken (instantiëren) van gegevens van bijvoorbeeld een metamodel naar een steeds lager niveau van ontsluiting maakt opbouw en controle van het datamodel mogelijk. Uiteraard kan een model ook bottom-up opgebouwd worden, waarbij steeds nieuwe metadata bronnen ontsloten worden. KNOWtory is ETL- of CASE-tool noch dbms, maar kan voor kleine hoeveelheden data de taken daarvan wel uitvoeren. Door testsets te gebruiken na consolidatie van meerdere bronnen, kan controle plaatsvinden op de uitgangspunten en uitwerkingen van de gevonden overeenkomsten en verschillen in de brongegevens. Hierdoor kunnen het gewenste eindresultaat en de processen om tot dit resultaat te komen in een vroeger stadium op elkaar worden afgestemd.

Een ander groot voordeel van deze manier van opslag is dat

De structuur van de aangeleverde data blijft behouden

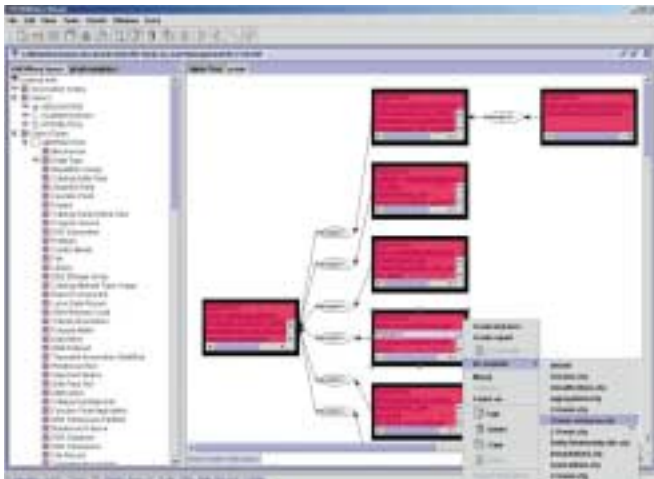
verschillende definities van eenzelfde gegeven kunnen bestaan. Daardoor kan op verschillende manieren tegen hetzelfde gegeven kan worden aangekeken. Ook kunnen meerdere modellen naast



FIGUUR 1: GRAFISCHE WEERGAVE VAN DE 'SUBJECT AREAS' IN EEN INFORMATIEMODEL VOOR EEN BANK, MET EEN CONTEXTMENU.

elkaar bestaan. Dat maakt het mogelijk verschillende ontwikkelmodellen op te bouwen en de resultaten te vergelijken.

Verder maakt KNOWtory de toegang tot de metadata eenvoudig. Via de verschillende modules, zoals Visual, Web of SDK, kunnen technische, functionele en systeemgebruikers bij de gegevens die zij nodig hebben om hun taken optimaal te kunnen uitvoeren. Rapportages kunnen zij zelf maken en standaard laten weergeven op bijvoorbeeld het intranet van de organisatie. Maar het is ook mogelijk via een publish and subscribe-faciliteit gebruikers te



FIGUUR 2: HYPERGRAPH VIEW VOOR HET NAVIGEREN IN GROTE MODELLEN.

informereren die alleen op de hoogte gehouden willen worden, maar geen behoefte hebben rechtstreekse toegang tot de metagegevens.

LIFECYCLE-SUPPORT

Traditionele metadatatools hebben voornamelijk meerwaarde aan het einde van de metadata-lifecycle, en dan met name voor technische metadata. Zodra het datawarehouse echt in productie gaat, krijgen de metagegevens hun functie in de warehouse-omgeving.

Met name de ETL-processen zijn vaak metadatagestuurd en worden bijgewerkt zodra de noodzaak ertoe bestaat. Aanpassingen in het warehouse zelf, maar ook in de eventuele datamarts op het warehouse worden al minder vaak compleet verwerkt in de tools en documentatie. Het is te omslachtig.

De diverse systemen zijn niet goed met elkaar geïntegreerd. Daardoor ontbreekt het totaaloverzicht. Vooral kleine wijzigingen worden al snel uitgevoerd, maar niet gedocumenteerd. Voor eindgebruikers zijn dit soort wijzigingen binnen afzienbare tijd niet meer inzichtelijk. Ook aanpassingen op de ETL-processen zijn voor deze gebruikers in de meeste gevallen niet toegankelijk. Toch is het gewenst dat een eindgebruiker weet waar zijn gegevens uit een rapport vandaan komen en wat er met deze data is gebeurd van de bronsystemen tot het eindresultaat op zijn rapport.

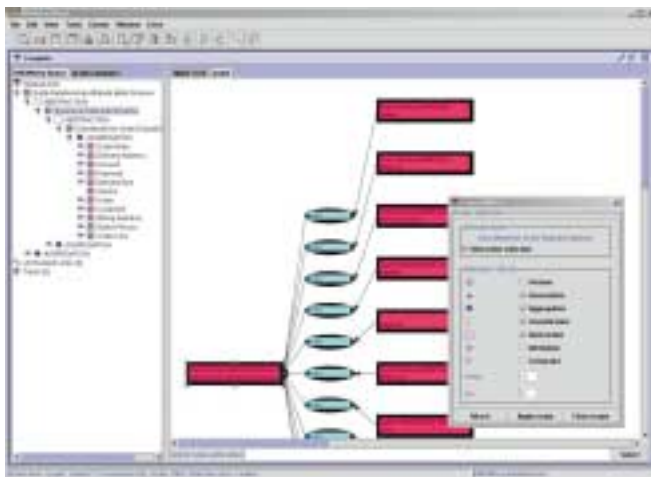
Bij grotere wijzigingen wordt het echter pas echt verwarrend. Oude legacy systemen zijn vaak ingelezen in het warehouse op aannamen en verwachtingen, omdat dit soort systemen veelal niet of slecht gedocumenteerd zijn en de informatie bij elkaar moet worden gesprokkeld. Zodra een dergelijk systeem wordt vervangen, heeft dit grote impact op het warehouse. Naast het feit dat er een duidelijke voor- en na situatie ontstaat, die voor de eindgebruiker inzichtelijk moet zijn en blijven, is het van groot belang voor de opbouw van het datawarehouse en de datamarts om te weten wat de consequenties zullen zijn van een wijziging van deze omvang. Dit inzicht is op de volgende manier te verkrijgen. Laat in het beginstadium van de ontwikkeling -of bij een standaard pakket de implementatie- van een nieuw systeem de metadata op een

Functionaliteit in 't kort

Data in KNOWtory niet volgens een vast stramen te worden ingelezen en opgeslagen en zijn daardoor flexibel te presenteren. Doordat ook de business-gegevens kunnen worden opgeslagen, zal er een veel grotere betrokkenheid en acceptatie van de eindgebruikers zijn, een belangrijke voorwaarde voor een succesvolle implementatie van een metadatatoeëpassing. Bovendien kan het tool direct gebruikt worden om de metadata in kaart te brengen (legacy en nieuwe metadata), waardoor snel resultaat getoond kan worden. KNOWtory is vanaf het begin van de metadata-lifecycle inzetbaar. De software is geschikt om een metadatatoeëpassing evolutionair neer te zetten, vanwege zijn schaalbaarheid en goede mogelijkheden voor uitbreiding met nieuwe metadata, bronnen en modellen. Op kleine schaal kunnen databasemodellen, ETL-programmatuur en dergelijke worden uitgetest, om te zien of de gegevens 'eruit zien' zoals werd verwacht. Naast rapportage- en analysefunctionaliteit, is voorzien in een mogelijkheid tot publish and subscribe voor gebruikers die wel op de hoogte gehouden willen worden, maar geen kennis van KNOWtory hebben of willen hebben.

gemakkelijke en toegankelijke manier aansluiten bij de gegevens in de al bestaande metabase. Ook het draaien van proef-runs is een optie om te kijken wat eventuele effecten zullen zijn. Zonodig snel bijsturen kan daardoor. Helaas ondersteunen de traditionele systemen deze faciliteiten voor *overview* en *impact* niet.

KNOWtory lijkt dit alles wel te kunnen. Al bij het begin van de bouw van bijvoorbeeld een datawarehouse valt snel resultaat te boeken op de metadata-ondersteuning, die in volgorde de snellere ontwikkeling van de datawarehouse-omgeving kan bevorderen. De metabase is makkelijk toegankelijk voor eindgebruikers. Zo raken zij eerder en vollediger betrokken bij de opzet van de metabase, maar ook bij die van het datawarehouse. De kans is bovendien groot dat zij betrokken blijven.



FIGUUR 3: ER-DATAMODEL MET 'DIAGRAMATIC DRILL VIEW' VAN EEN PROPERTY SHEET.

Uiteindelijk gelden deze principes natuurlijk voor elke omgeving, van een operationeel databasesysteem tot de speciaal aan de frontend ontwikkelde applicaties.

KNOWTORY-TOOLSET

De KNOWtory-toolset bestaat uit de volgende modules:

- Enterprise Server, de basis van de toolset, die zorgt voor de schaalbaarheid en de integriteit van de data;
- Web, een gebruikersinterface voor interactieve analyses en het ontwikkelen van rapportages. Verzorgt de mogelijkheid tot toegang en bewerking van de opgeslagen gegevens;
- Visual, met alle mogelijkheden van KNOWtory/Web aangevuld met standaard methodologieën als UML;
- Administrator, een add-in voor Visual of Web, die zorgt voor de administratieve taken als autorisaties en registratie;
- Lifecycle, add-in voor de Web- of Visual-module, die bestaat uit een set van services, waaronder het agenderen van routinematige taken op basis van gebruik of content triggers;
- SDK, met een verscheidenheid aan API's. Ontwikkelaars kun-

KNOWtory kan over het algemeen goed met de complexe metadatatproblematiek uit de voeten

nen de KNOWtory-omgeving naar eigen behoefte verder uitbreiden;

- XML batch processor, waarin zware batch processen kunnen worden gepland.

CONCLUSIE

Voor iedereen die de frustratie kent van het zoeken naar gegevens en alle informatie omtrent die gegevens, en met die reden wil werken aan de opzet van een metabase, zal dit tool een grote stap in de goede richting kunnen zijn. KNOWtory is zeker de moeite van een evaluatie waard. Dat geldt nog sterker voor organisaties die de afgelopen tijd daadwerkelijk zijn begonnen met de bouw van een metabase en daarbij ontdekten dat dit proces nog steeds niet goed wordt ondersteund door de systemen in de markt.

Aangezien het product koud op de markt is, zal het zich in de nabije toekomst verder moeten gaan bewijzen. KNOWtory is nog niet grootschalig toegepast. Ook zitten hier en daar nog wat verbeterpunten, die in volgende versies verholpen zullen gaan worden. ●

Sandra Gerrits (sandra.gerrits@eniklas.com) is consultant bij eNiklas Netherlands, een in business intelligence gespecialiseerde ICT-dienstverlener.