

# Kennisintensieve Systemen in Onderzoek

**Tamme van der Wal**

Manager Group Software Engineering

DLO Staring Centrum, Postbus 125, 6700 AC Wageningen

telefoon (0317) 47 42 31/47 42 00; telefax (0317) 42 48 12

e-mail: T.vanderwal@sc.dlo.nl

Kennis is een abstract begrip, eigenlijk van filosofische aard. Om kennis te onderscheiden van data en informatie is alleen al een moeilijke zaak. Een prettig werkbaar definitie is te vinden bij 'Center for Enterprises Systems'. Data, Informatie en Kennis bevinden zich in een continuüm, met data aan de ene en kennis aan de andere kant: Data zijn ongecorrleerde feiten en metingen; Informatie is het resultaat van het correleren en organiseren van data; Kennis stelt de 'gebruiker' in staat om uitspraken te doen over nieuwe situaties [Tjaden].

Kennis kan worden verkregen en kennis kan worden gebruikt. In het kennisontwikkelingsmodel van Nonaka & Takeuchi (o.a. in [Smit]) wordt de wisselwerking tussen kennis gebruiken en kennis verkrijgen beschreven als een continu proces. Tjaden betoogt dat zodra bepaalde kennis gecreëerd is, het vele malen hergebruikt kan worden. Echter, bepaalde kennis wordt overbodig wanneer het overtroffen wordt door nieuwe kennis.

onderzoek is de vastlegging en ontsluiting van bestaande data, informatie en kennis noodzakelijk.

Beide typen geautomatiseerde systemen worden ingezet bij het doen van uitspraken over een toepassingsgebied. Het functioneert echter veelal als tweetraps raket; Het proces 'uitspraken doen over' wordt ondersteund door het model, en het model behoeft de repository. Hiermee kunnen we repositories classificeren als data-opslag- of informatiesystemen. De werkelijke kennis ligt besloten in het simulatiemodel.

De mate waarin simulatiemodellen het onderzoek ondersteunen is nogal verschillend. De effectiviteit van een simulatiemodel kan daarom getypeerd worden aan de hand van drie variabelen: Complexiteit, Integratie en Dynamiek.

Met de complexiteit van een kennisysteem wordt hier bedoeld het aantal stappen dat door het systeem bestreken wordt. In andere woorden: complexe systemen vereisen minder tussenkomst van de gebruiker dan simpele systemen. Als gevolg daarvan kunnen we stellen dat complexe systemen minder specifieke domeinkennis van de gebruiker vereisen dan simpele modellen. Veel huidige modellen kunnen in deze terminologie bestempeld worden als simpele systemen, daar alleen met behulp van kennis van het toepassingsgebied zinnige simulaties verkregen kunnen worden (denk hierbij bijvoor-

Er is nu een beeld geschetst van kennis. De vraag blijft nog wat kennisintensief dan betekent. Kenniscentrum CIBIT biedt uitkomst. In hun folder voor de MSc cursus Information and Knowledge Technology omschrijven ze kennisintensieve processen als processen 'die waarde toevoegen aan gegevens, op grond van kennis over het toepassingsgebied'. Voorbeelden van dergelijke processen zijn: besluiten, adviseren, plannen, analyseren, beoordelen, herkennen en oplossen. Volgens CIBIT zijn kennisintensieve systemen dan systemen die ingezet worden bij kennisintensieve processen.

Op grond van het bovenstaande luidt dan mijn volledige definitie van een kennisintensief systeem: Een systeem dat door verwerking van kennis behulpzaam is bij het doen van uitspraken in een bepaald toepassingsgebied.

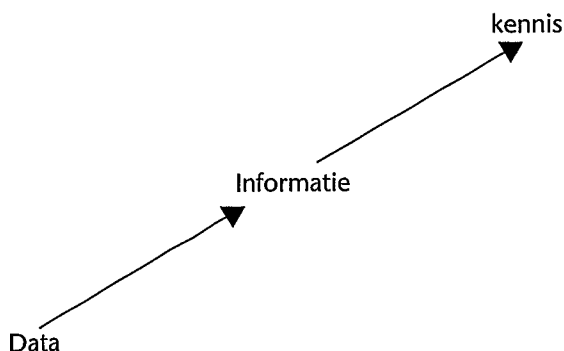
matiseerde systemen. In het kader van dit artikel wil ik graag twee typen geautomatiseerde systemen naar voren halen:

- Modellen: Een belangrijk stuk gereedschap voor de huidige onderzoeker is een simulatiemodel. Met behulp van formele definities en structuren maakt de onderzoeker een afspiegeling van (een deel van) de werkelijkheid. Door middel van simulatie wordt de werkelijkheid 'nagespeeld'. Het simulatiemodel wordt ook wel als virtueel laboratorium gebruikt.

- Repositories: Een tweede belangrijke categorie kennisintensieve systemen zijn repositories. In een ruime definitie verstaan we hieronder databases (ook GIS bestanden) en bijbehorende zoeksystemen. Voor

## Kennisintensieve systemen en onderzoek

Het toepassingsgerichte onderzoek bij DLO wordt onder andere gekenmerkt door het doen van uitspraken over het desbetreffende vakgebied. Dit soort werk wordt in veel gevallen ondersteund door geauto-



beeld aan data preprocessing). In termen van effectiviteit kan men streven naar complexere systemen, waardoor ook gebruikers met minder domeinexpertise zinnige uitkomsten kunnen genereren.

Integratie doelt op de mate waarin het systeem het proces ondersteunt, dan wel adequaat ingezet kan worden. Daar valt ook onder de mate van integratie met andere systemen (die deel uitmaken van een proces). Enkele uitzonderingen daargelaten kunnen we stellen dat het merendeel van de simulatiemodellen ingericht is voor speciale vragen en processen. Het is meestal specifiek ontwikkeld en daarmee ver geïntegreerd met dat proces. Hetzelfde simulatiemodel kan daarmee nog niet voor andere processen effectief ingezet worden. Daarmee zijn we aangeland bij de derde variabele namelijk de Dynamiek. Hiermee wordt de mate waarin een systeem aangepast kan worden aangegeven. Ons specifiek ingerichte instrumentarium is veelal niet dynamisch.

Veel simulatiemodellen zijn nog geen effectief kennisintensief systeem. Immers een effectief systeem vereist weinig specialistische kennis, is in hoge mate geïntegreerd met de processen die het moet ondersteunen en is eenvoudig aan te passen aan verandering in de processen.

## De vraag naar effectieve kennissystemen

Het onderzoek van DLO heeft de afgelopen jaren een aantal ontwikkelingen doorgevoerd. Allereerst stellen beleidsmakers en andere klanten vragen van een meer geïntegreerd karakter. Onderzoeksgroepen kun-

nen dergelijke projecten niet meer alleen af en projecten worden daarmee multidisciplinair en zelfs instituut overstijgend. Ook is de problematiek elke keer weer anders en de vragen vanuit beleid veranderen snel. Verder wordt bij het onderzoek meer gebruik gemaakt van simulatiemodellen. Dit wordt met name veroorzaakt door de toegenomen reken capaciteit en door de steeds betere beschikbaarheid en ontsluiting van digitale bestanden. En tot slot vragen klanten meer dan vroeger om gereedschap om zelf studies uit te voeren.

Deze ontwikkelingen leggen een grote druk op het bestaande modelinstrumentarium. Steeds nieuwe vragen leiden tot wijzigingen, uitbreidingen en soms zelfs tot compleet redesign van bestaande systemen. Maar de toenemende snelheid waarmee systemen moeten veranderen binnen beperkte budgetten vereist een andere aanpak.

Een verdergaande modularisering lijkt de enige oplossing. Door de juiste componenten samen te voegen tot een systeem krijgt de klant maatwerk. Door hergebruik worden individuele componenten goedkoper (in bouw en onderhoud). Het streven naar maatwerk door modularisatie wordt ook wel de 'Economies of Scope' genoemd [Smit]. Economies of Scope streeft naar het inzetten van vaste middelen voor het leveren van maatwerk diensten.

Naast een assortiment componenten om mee te combineren, is voor een effectief instrumentarium een goede architectuur onontbeerlijk. De samenhang van componenten, de wijze waarop ze geacht worden te interacteren en de wijze waarop de rollen en verantwoordelijkheden tussen software-

componenten verdeeld worden, dient zorgvuldig doordacht en vastgelegd te worden in een systeemarchitectuur. Immers uw aanpak gaat uw huis ook niet bouwen zonder tekening en zomaar wat stenen stapelen en kozijnen plaatsen.

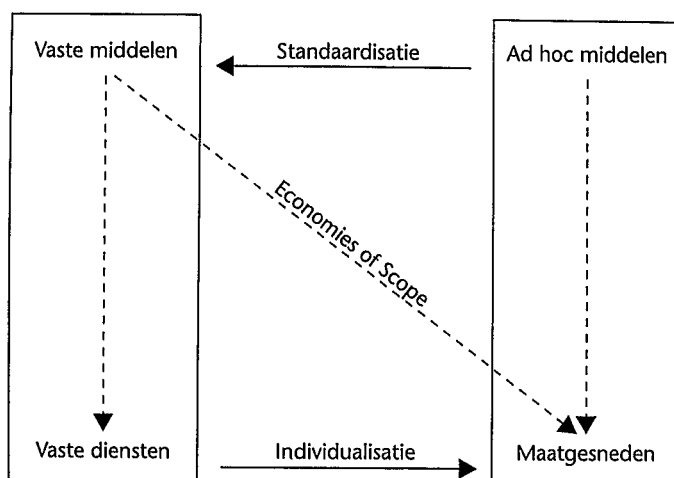
## Ontwikkeling van kennisintensieve systemen

Software speelt een belangrijke rol in het onderzoek van DLO Staring Centrum. In 1996 is daar de GROUP SOFTWARE ENGINEERING (GSE) opgericht. GSE heeft als missie om de ontwikkeling van software ten behoeve van het onderzoek te professionaliseren. GSE heeft een lijn uitgestippeld om het geïntegreerd modelinstrumentarium voor de komende vijf tot tien jaar te ontwikkelen. Dit wordt nagestreefd door de inzet van objectgeoriënteerde methoden en nieuwe technieken om tot systeem beschrijving en requirements specificaties te komen.

### Domein Analyse

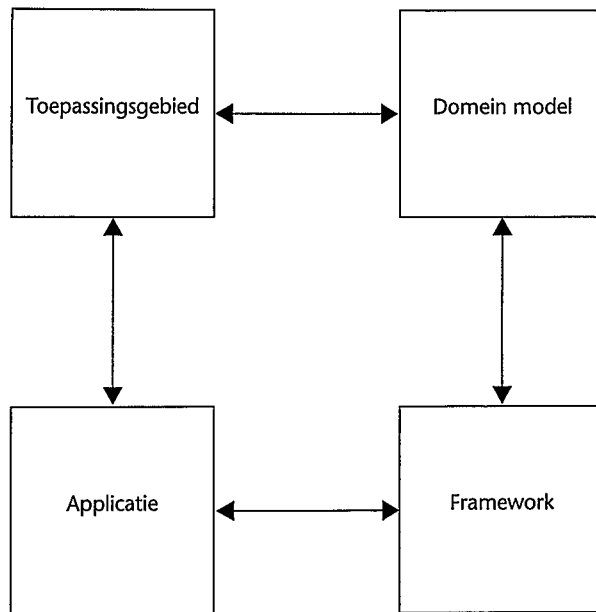
Om tot betere integratie en afstemming van (deel)modellen te komen wordt bij de systeemontwikkeling niet primair van de functionele specificaties uitgegaan maar van het vakdomein waartoe een systeem behoort. Dit vakdomein wordt in kaart gebracht in de zgn. Domein Analyse. Die wordt uitgevoerd door software engineers en domein deskundigen. Met behulp van de Unified Modelling Language (UML) wordt het desbetreffende domein in kaart gebracht. Het Domein Model vormt een prima communicatiemedium. Het Domein Model beschrijft het toepassingsgebied van een simulatiemodel. Door dit zo getrouw mogelijk aan de werkelijkheid te maken wordt het systeem flexibeler. Immers alle conceptuele bouwblokken uit de werkelijkheid hebben hun plaats.

Voor enkele domeinen heeft onder begeleiding van GSE zo'n analyse plaatsgevonden. Enkele voorbeelden van deze domeinen zijn 'Integraal Waterbeheer', 'Precisielandbouw'



Figuur 1 – Economies of Scope (Smit & Bruin, 1998).

**Figuur 2 – Samenhang tussen de verschillende begrippen.**



en 'Bestrijdingsmiddelen'. Hiervoor wordt gebruik gemaakt van participatieve technieken. Het eigenlijke Domein Model wordt dan ook door onderzoekers opgezet. Software Engineers zijn hierin uitsluitend faciliterend. Het blijkt dat UML een krachtig hulpmiddel is om domeinen in kaart te brengen. Deze notatiewijze ondersteunt het in kaart brengen van de verschillende domeincomponenten en hun onderlinge relaties. Aan componenten worden specifieke functies en verantwoordelijkheden toegewezen. Hierdoor ontstaat een helder en consistent beeld van het domein. Een ander voordeel van het gebruik van UML is dat het domeinmodel en de verdere ontwikkeling daarvan bijgehouden kan worden tijdens de design-en bouwfasen van een project.

### Frameworks

Daar waar meerdere applicaties met hetzelfde domein te maken hebben, kan het Domein Model bijdragen in het herkennen van herbruikbare delen. Verder zal de samenhang van de componenten (relaties, rollen en verantwoordelijkheden) niet veel veranderen. Een specificatie van de samenhang van componenten in een bepaald toepassingsgebied noemen we een Framework. Daarnaast omvat een framework ook instructies over hoe het gebruikt kan worden. In een framework worden "standaard" oplossingen aangedragen voor "standaard" problemen. Dit kunnen standaardoplossingen zijn in het toepassingsdomein (bijv. een bepaalde berekeningswijze voor Verdamping) en oplossingen in het applicatiedomein (bijv. een bepaalde grafische weergave van een getallenreeks). Hierdoor maken we onderscheid in een applicatieframework en een domeinframework. In verschillende projecten wordt momenteel gewerkt aan het ontwikkelen van frameworks.

Applicaties tenslotte worden ontwikkeld

met gebruik van het framework. Applicaties worden ingezet ter ondersteuning van problemen in het toepassingsgebied. De standaard of generieke oplossingen van het framework kunnen direct gebruikt worden of eventueel aangepast of uitgebreid worden naar behoefte.

In figuur 2 geef ik de verschillende begrippen uit het bovenstaande schematisch weer. De werkwijze die we bij GSE hanteren voor het ontwikkelen van kennisintensieve systemen begint bij het toepassingsgebied en gaat vervolgens met de klok mee. Het omgekeerde is ook mogelijk. Dan wordt alles opgehangen aan de bestaande applicaties. Dit is echter een minder wenselijke route omdat het een onvolledig of zelfs een onjuist beeld kan geven.

### Ontwikkelen voor de toekomst

Technisch wetenschappelijke software is in de afgelopen 20 jaar hoofdzakelijk ontwikkeld voor specifieke doeleinden en voldoet niet aan een aantal kwaliteitskenmerken zoals koppelbaarheid, uitbreidbaarheid, naspeurbaarheid en robuustheid. Veel huidige systemen zijn niet of nauwelijks toegerust voor ondersteuning van het onderzoek in de komende tien jaar. Vernieuwing is

onontbeerlijk om de komende jaren het onderzoek adequaat te ondersteunen met kennisintensieve systemen. Door systemen te baseren op de toepassingswereld en door modularisatie kunnen applicaties gemaakt worden die aan de eisen van nu en de (nabije) toekomst voldoen.

DLO heeft met GSE een innovatieve dienstverlener op het gebied van de informatietechnologie. Met nieuwe methoden en technieken uit die IT wereld draagt GSE bij aan het verbeteren van het onderzoek.

### Verantwoording

Group Software Engineering is een samenwerkingsverband tussen DLO Staring Centrum en Q-Ray Agrimathica en werkt in projectverband aan opdrachten voor Staring Centrum, DLO en externe opdrachtgevers. Samen met Software Engineering Research Centre (SERC) werkt GSE aan de ontwikkeling van een framework voor het integraal waterbeheer.

### Hergebruikte kennis

Tjaden G.S., *Measuring The Information Age Business*, Center for Enterprise Systems, 1995.

Smit W.W.M. & W.H. de Bruin, *Excelleren of Creperen*, Uitgeverij Contact, Amsterdam, 1998. @