

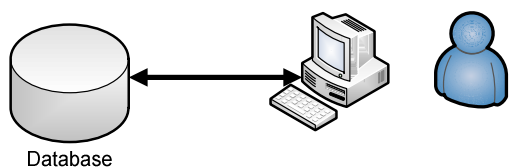
Activiteiten Lectoraat Data Architectures en Metadata Management

De Hogeschool van Arnhem en Nijmegen (HAN University of Applied Sciences) speelt al 15 jaar een landelijke voortrekkersrol op het gebied van innovatieve informatiesysteemontwikkeling, gebaseerd op conceptuele modellering, geautomatiseerde modeltransformaties en volledig metadata driven toolsupport. Met de instelling van het lectoraat Data Architectures en Metadata Management is een structureel platform gecreëerd voor verder onderzoek en voor het ten dienste stellen van de opgebouwde expertise aan het hoger onderwijs en het bedrijfsleven in binnen- en buitenland.

Hieronder volgt een overzicht van de onderzoeksgebieden en activiteiten van het lectoraat.

Informatiesystemen

Een groot deel van de computersystemen zijn zogenaamde data-intensieve informatiesystemen waarbij gebruikers gestructureerde informatie opvragen en bewerken. Alle gegevens worden bewaard in een grote gegevensbank (database) en gebruikers hebben de beschikking over een programma (applicatie) om met menu's en allerlei schermen (forms) en formulieren (reports) op een gebruikersvriendelijke manier de database te benaderen.

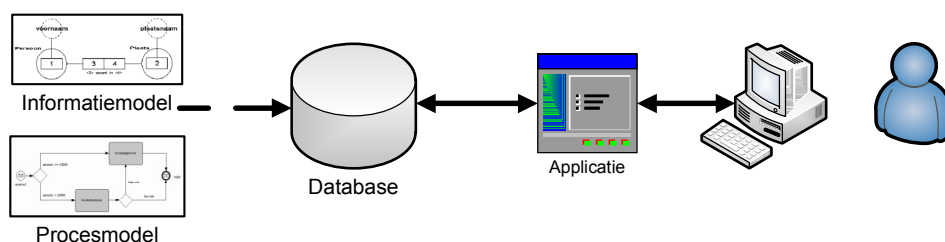


Voorbeelden zijn informatiesystemen voor banken, verzekeringsmaatschappijen, ziekenhuizen, vluchtreserveringssystemen en studievoortgangssystemen van hogescholen en universiteiten.

Het gehele systeemontwikkelingstraject van data-intensieve systemen is het aandachtgebied van het lectoraat Data Architectures & Metadata Management.

ICA-opleidingen

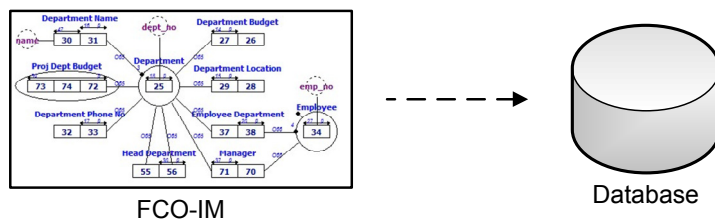
In het ICA bachelor-onderwijs wordt een en ander gedoceerd in de semestercurcus 'Develop an Information System': de behoeften van de klant moeten worden geanalyseerd, wat leidt tot een informatiemodel dat wordt omgezet in een redundantievrije database. Tenslotte moeten de diverse schermen van de applicatie worden ontwikkeld m.b.v procesmodellen en andere specificaties. Voorts is er vanuit het lectoraat een Engelstalige minor 'Behind the Screens' ontwikkeld en is een minor 'Business Intelligence' in ontwikkeling. De nog beduidend bredere en diepergravender Engelstalige ICA-masteropleiding 'Information Systems Development' wordt al vele jaren uitgevoerd en werd in 2008 weer positief geaccrediteerd.



Methoden en technieken

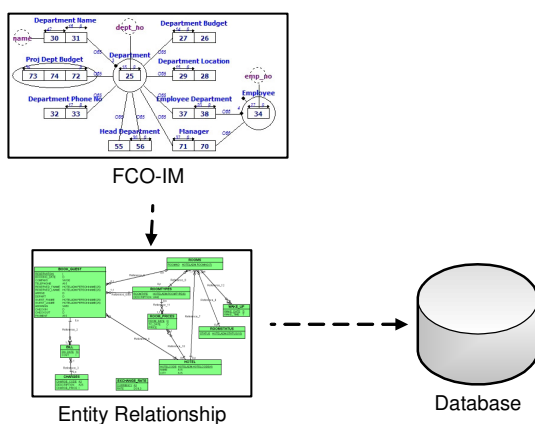
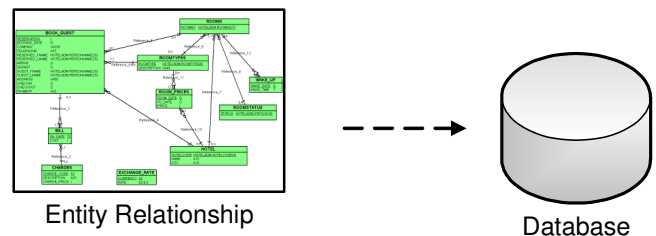
Voor het ontwerpen van een informatiemodel bestaan verschillende methoden. Binnen de onderzoeksgroep is gekozen voor het verder ontwikkelen van Fact Oriented Methods (FOM), die natuurlijke taal gebruiken om op basis van concrete voorbeelden conceptuele informatiemodellen te maken. Dit heeft geleid tot FCO-IM (Fully Communication Oriented Information Modeling), een nieuwe FOM-methode waarbij ook het resultaat beter gevalideerd kan worden met de opdrachtgever/gebruikers dan het geval was bij oudere FOM-methoden als NIAM en ORM.

Vervolgens kan een databasestructuur worden verkregen door een transformatie (conversie) van het gevalideerde FCO-IM informatiemodel, waarbij wordt gegarandeerd dat alle data slechts éénmaal in de (genormaliseerde) database wordt vastgelegd. Tegenstrijdigheden worden daardoor voorkomen en ontwikkel- en onderhoudskosten worden verlaagd.



Een belangrijk uitgangspunt voor verder onderzoek is dat ook regels over de applicatie, zoals wie mag wat, slechts één maal in het systeem worden vastgelegd. Doorontwikkeling van FCO-IM op dit aspect is een speerpunt van het lectoraat.

Andere technieken voor datamodellering zijn o.a. Entity-Relationship Modeling (ERM) en UML Class Diagramming, die - althans qua datamodellering - redelijk vergelijkbaar zijn. Deze technieken zijn weliswaar onder IT-ers meer vertrouwd (worden overal onderwezen), maar zijn conceptueel (validatie door gebruikers is niet- of nauwelijks mogelijk) en logisch (een redundantievrije database kan niet worden gegarandeerd) beduidend minder krachtig.



Uit onderzoek binnen het lectoraat is gebleken dat ook ERM modellen en UML class-diagrammen verkregen kunnen worden door transformaties vanuit een FCO-IM model. Deze transformatiemethoden worden toegepast in de bacheloropleiding en diepgaand bestudeerd in de masteropleiding.

Er is onderzoek gedaan naar de integratie van proces- en gegevensmodellering door master-studenten in afstudeerprojecten, maar hier is nog het nodige te doen.

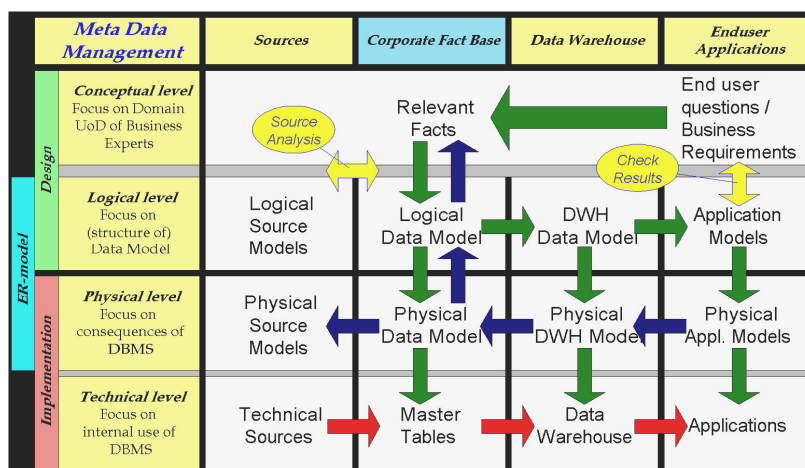
Datawarehousing en Business Intelligence

Naast het opvragen van eenvoudige rapportages bestaat er in de praktijk vaak behoefte aan management rapportage. Deze overzichten bevatten geaggregeerde gegevens waarvoor vaak veel berekeningen nodig zijn. Voorts moet de rapportage flexibel zijn ten aanzien van het weergeven van meer (drill down) of minder (drill up) details. Men spreekt van Business Intelligence.

Het doorrekenen ten behoeve van deze rapportages is te kostbaar voor de productiesystemen (transactionele systemen). Mede daarom worden dagelijks/wekelijks/maandelijks kopieën van de database gemaakt en opgeslagen in een andere databasestructuur; men spreekt men van een datawarehouse, of kleiner en dedicated voor bijvoorbeeld een enkele afdeling, van datamarts. Om de snelheid van berekeningen en de flexibiliteit te verhogen, zijn speciale modelleertechnieken voor datawarehouses en met name voor datamarts ontwikkeld. De meest populaire is de multidimensionele modelleertechniek, waarbij zogenaamde sterren en kubussen een rol spelen. Voor het realiseren van de applicaties voor het benaderen van een datawarehouse of datamart worden speciale front-end tools (zg OLAP tools) gebruikt, maar tegenwoordig kom je ook een heel eind met speciaal ingerichte Excel-bestanden. In een datawarehouse-omgeving bestaat meer en meer aandacht voor het mee-analyseren van historische informatie en zelfs van wijzigingen op die historie (historie van historie).

Metadata Frame

In de onderzoeksgroep is veel energie gestoken in het specificeren van een algoritme voor de transformatie van een FCO-IM model naar dimensionele modellen (families van sterren), zoals tegenwoordig in gebruik in Business Intelligence toepassingen. Tevens is een FCO-IM based framework ontwikkeld voor het ontwerpen en realiseren van informatiesystemen, zowel voor productiesystemen als voor business intelligence systemen. Dit framework is door Atos Origin integraal overgenomen als Metadata Frame.



Gereedschappen voor modellering

Voor data- en procesmodelleringstechnieken bestaan allerlei gereedschappen (tools), die zijn gebaseerd op een gestructureerde bibliotheek (repository) waarin ontwerpgegevens (metadata) kunnen worden opgeslagen en een diagrammatische notatie van die vastgelegde ontwerpgegevens. Voor FCO-IM zijn toolrepositories ontworpen en twee FCO-IM tools gespecificeerd: een FCO-IM modelleringstool en een toolset voor de gewenste model-to-model transformaties. Van het eerste tool zijn eerst door afstudeerstudenten van de eigen en andere hogescholen opeenvolgende proof-of-concept realisaties gemaakt, waarna door 3 softwarebedrijven hiervoor commerciële tools zijn gebouwd: Commit (door TLO Holland Controls), Infagon (door Ascaris Software) en CaseTalk (door BCP Software), welke in

toenemende voldoen aan de specificaties. Met name CaseTalk (www.CaseTalk.com) wordt in het hoger onderwijs en in diverse praktijkprojecten met succes ingezet.

Binnen de onderzoeksgroep zijn in master-niveau afstudeerprojecten en door R&D-medewerkers van het lectoraat algoritmen gespecificeerd en proof-of-concept tools ontwikkeld voor model-to-model transformaties vanuit FCO-IM: ERBridge, UMLBridge en StarBridge. Deze losse Bridge tools zijn onder één generieke architectuur verenigd in het tool FCO-IM Bridge, dat naadloos aansluit op CaseTalk. Deze gecombineerde tooling (CaseTalk + FCO-IM Bridge) is door medewerkers van Atos Origin in de context van het Metadata Frame met succes toegepast in diverse grote datawarehousing-projecten, ondermeer bij KLM, Rabobank, Erasmus Medisch Centrum en recent het CBS. Het project in het Erasmus MC heeft de Computable Award 2007 gewonnen als “het meest innovatieve project in de Zorg”.

Gereedschappen voor applicatieontwikkeling

Bij conventionele applicatieontwikkeling worden schermen ‘met de hand’ gebouwd. Binnen het lectoraat is een generatietechniek ontwikkeld, waardoor deze fase van het hele proces enorm versneld kan worden. Natuurlijk moet het originele informatiemodel uitgebreid worden om vast te leggen welke gebruikers welke gegevens mogen zien dan wel wijzigen. Na een proof-of-concept realisatie van een Windows-applicatiegenerator, werd binnen het lectoraat de krachtige en innovatieve Web-applicatiegenerator IMAGine ontwikkeld. IMAGine past schermen van de gebruiker automatisch aan aan de gemodelleerde rechten. Met behulp van IMAGine is voor de HAN op deze manier onder meer het studievolsysteem HANSIS ontwikkeld. Deze wijze van ontwikkeling sluit naadloos aan op de FOM informatiemodelleermethoden.

Cursussen en trainingen

De besproken methoden en technieken voor informatie- en procesmodellering, worden niet alleen onderwezen in de bacheloropleiding en in de masteropleiding, maar tevens voor externe bedrijven. Ook wordt dan diep ingegaan op de model-to-model transformaties en het Metadata Frame. Naast aan de door ons ontwikkelde generatie op basis van directe interpretatie van de bedrijfsregels, wordt in het onderwijs ook aandacht besteed aan het genereren van applicatiecode.

Ondermeer is een 4-daagse in-company training ‘Hands on FCO-IM’ meerdere malen uitgevoerd bij Atos Origin, waarbij in totaal ondermeer 150 medewerkers zijn getraind. Vergelijkbare meerdaagse trainingen zijn verzorgd voor IT-medewerkers van KLM, Rabobank, Gemeente Apeldoorn en andere bedrijven en instellingen en voor teams van hogescholen. Driemaal is een 2-weekse training ‘Information Modeling and Database Design with FCO-IM’ verzorgd voor het Amerikaanse bedrijf Sandia National Laboratories. Door de ontwikkelgroep van IMAGine zijn ook daarvoor trainingen verzorgd.

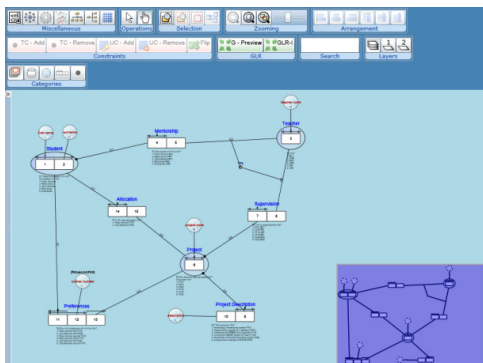
Op uitnodiging zijn in het verleden gastcolleges en seminars verzorgd in Nederlandse (Nijmegen, Delft) en buitenlandse universiteiten (Costa Rica, Marokko, Duitsland, UK, Albanië, Vietnam, Indonesië, Maleisië). Met China zijn tevens contacten op dit gebied.

Publicaties

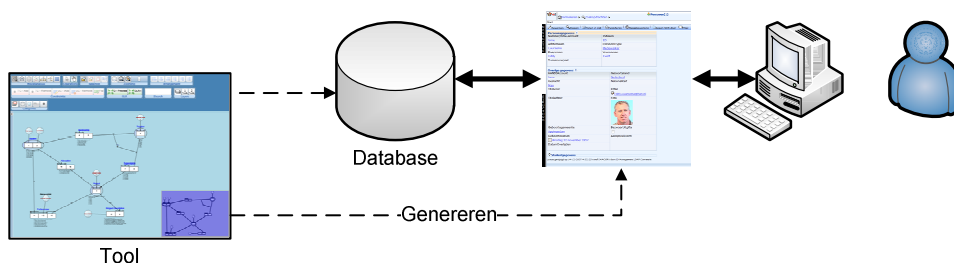
Naast de boeken 'Volledig Communicatiegeoriënteerde Informatiemodellering (FCO-IM)' en 'Volledig Communicatiegeoriënteerde Informatiemodellering, Opgaven en casussen' (beide ook als pdf-files in het Engels beschikbaar), zijn in de afgelopen jaren diverse publicaties verschenen die zijn gepubliceerd in Tinfon, Database Magazine, the Journal for Conceptual Modeling en op internationale conferenties. Zie daarvoor de FCO-IM website: www.FCO-IM.nl.

Recente toolontwikkelingen

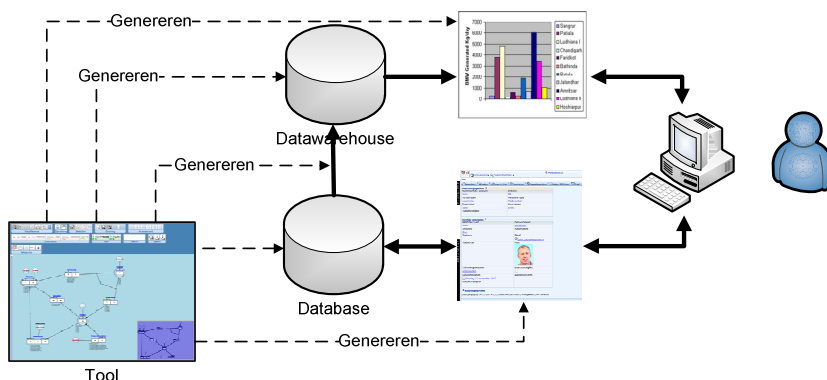
Binnen het lectoraat is inmiddels een universeel basisgereedschap in ontwikkeling genaamd Graphity, waarmee allerlei modelleringstechnieken (informatie, proces, organisatie) kunnen worden ondersteund. Met dit gereedschap wordt thans een Web-gebaseerd FCO-IM modellertool met grafische presentatie in FCO-IM, ERM, UML, Relationeel, etc. ontwikkeld.



Ook wordt onderzoek gedaan om vanuit een conceptueel FCO-IM informatiemodel naast een database ook een Web-applicatie te genereren. De eerste resultaten zijn ondergebracht in een product genaamd FCO-IMagine.



Vervolgens moeten ook een corresponderend Dimensioneel datawarehouse, de nachtelijke kopieën daarheen en de gewenste OLAP Excel-sheets grotendeels automatisch worden gegenereerd.



Stap voor stap zullen de benodigde transformaties diagrammatisch met Graphity worden ondersteund.

Deelprojecten

In het komende jaar (jaren) wordt er in het lectoraat gewerkt aan de volgende deelprojecten:

- Project Advances in FCO-IM: verder uitbouwen van de theorie van FCO-IM en daarover wetenschappelijk publiceren.
- Publiceren over diverse ontwikkelingen in Database Magazine.
- Verder uitbouwen van de applicatiegenerator IMAGine.
- Verder uitbouwen van de toolgenerator Graphity.
- Realiseren van een Web-gebaseerd FCO-IM modelleertool.
- Realiseren van een Web-gebaseerde generatie van dimensioneel datawarehouses uit een FCO-IM model, gebaseerd op eerder theoretisch/prototype werk StarBridge
- Onderzoek naar de modelleertechniek DataVault voor historie in datawarehouses en naar FCO-IM naar DataVault conversie.
- Onderzoek naar procesmodellering in combinatie met data/applicatiemodellering.
- Beheer en onderhoud HANSIS.
- Schrijven van nieuwe hoofdstukken voor een nieuwe herziene en uitgebreide druk van het FCO-IM boek of een tweede FCO-IM boek.

Nadere informatie

Hogeschool van Arnhem en Nijmegen (HAN), www.han.nl.

Informatica Communicatie Academie (ICA), www.ica.han.nl.

Lectoraat Data Architectures & Metadata Management, www.han.nl/start/graduate-school/onderzoek/lectoraten-kenniskringen/data-architectures-metadata-management.

Bezoekadres: HAN campus Arnhem, gebouw faculteit Techniek, Ruitenberglaan 26, kamer E1.03

Postadres: Postbus 2217, 6802 CE Arnhem

Secretariaat: Lia Venhof (e-mail: lia.venhof@han.nl, tel: 026-3658152, fax: 026-3645066)

Lector: drs. Guido Bakema (guido.bakema@han.nl)

Associate lector: ir. Eddy Luursema (eddy.luursema@han.nl)

Program manager masteropleiding: Chris Scholten MSc (chris.scholten@han.nl)

Promotieonderzoek Advances in FCO-IM: Jan Pieter Zwart (janpieter.zwart@han.nl)

FCO-IM website: www.FCO-IM.nl