

Ontsluiting van Geo-informatie via netwerken

Plan van aanpak

Drs. M.E. de Vries

Summary:

This report describes the context and the research goals of the PhD project 'Building blocks for geographic information exchange via Internet and mobile networks'. The project is financed by the section GIS-technology of the Geodesy department of TU Delft, by the Provincie Gelderland and by DIOC BETADE.

Purpose of the PhD research is the design of an open (=non proprietary) distributed architecture for the exchange of geographic information, where data and services are components that can be used in a 'plug-and-play' manner.

Sub goals are: decomposition of functionality into smaller building blocks, comparison of different technical implementations (configurations), building of prototypes, formulation of - maybe conflicting - evaluation criteria for quality of both architecture and components (open, extensible, replaceable, generic etc.).

Geographic information plays an increasingly important role in our society. In many situations reliable and up-to-date spatial information is crucial, e.g. in disaster response systems. In other cases geographic information is an important input for decision-making processes of governments (urban planning, nature conservation), businesses (market exploration, location analysis) or individuals (tourist information, real estate prices).

Internet has created new possibilities for the use and exchange of spatial information. Geo-data does not have to be copied anymore on CD's or tape, but could in principle be distributed by Internet technology. This has many advantages. 'Copying around' on physical media (CD's or tape) will lead to an unclear mix of versions of the data, to data inconsistencies and to information that is not up-to-date. On the other hand: when one leaves the data where it is produced and maintained ('data at the source') data quality will improve, users (business organizations and government agencies) will have access to up-to-date information when and where they need it, users can share information in stead of having to (re) produce necessary data themselves. In other words: one will have the benefits of a (Global) Spatial Information Infrastructure.

However, one of the main bottlenecks for the implementation of an open (=non proprietary) distributed architecture for the exchange of spatial information is the heterogeneous nature of geographic data, both in a technical sense (data format) and in a semantic sense (data model and meaning). Data exchange almost always means data conversion from one model or format into another. Interoperability (of data and software) is therefore one of the issues that needs to be resolved in order to achieve a (global, national or regional) spatial information infrastructure.

The main elements and the methodology of the research will be:

- 1. design of the overall architecture: decomposition of functionality at different levels of granularity (UML)*
- 2. interoperability: issues, existing initiatives (OpenGIS, ISO and W3C standardization), possible solutions (theory and prototype testing)*
- 3. existing 'proprietary' Internet GIS software: is it possible to 'extend' these legacy systems in such a way that they can play a part in an open data exchange framework (tests)*
- 4. configurations: component framework technologies, implementation choices (server or client side processing etc.) (theory and prototype testing)*
- 5. (iterative) testing of possible solutions in prototype implementations (case studies)*

ISSN: 1569-0245

© 2002 Sectie GIS technologie

TU Delft, Faculteit der Civiele Techniek en Geowetenschappen

Thijssseweg 11, 2629 JA Delft

Tel: 015-2784548; Fax 015-2782745; E-mail: gist@tudelft.nl

Niets uit deze uitgave mag worden verveelvoudigd en/ of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de sectie GIS technologie

De sectie GIS technologie aanvaardt geen aansprakelijkheid voor eventuele schade voortvloeiend uit het gebruik van de resultaten van dit onderzoek of de toepassing van adviezen.

Inhoud

1. Context van het promotie-onderzoek	3
1.1 'Building blocks' en software-componenten	3
1.2 Software-componenten en UML	4
1.3 Geo-informatie en netwerken	4
1.4 OpenGIS en interoperabiliteit	5
1.5 Communicatie-raamwerken	6
2. Doel van het promotie-onderzoek	7
3. Opzet en fasering	8
3.1 Fase 1 (eerste jaar)	8
3.2 Fase 2 (tweede en derde jaar)	8
3.3 Fase 3 (vierde jaar)	8
4. De case studies (prototypes)	9
5. Samenwerking met externe organisaties	12
6. Organisatie van het onderzoek	13
6.1 Opleiding en training	13
6.2 Begeleiding	13
6.3 Ander inhoudelijk overleg	13
7. Benodigde middelen	14
7.1 Software	14
7.2 Kosten congresbezoek	14
7.3 Kosten bezoek aan binnenlandse en buitenlandse instanties en onderzoeksinstellingen	14
7.4 Kosten gespecialiseerde opleiding en training	14
Resumé	15
Literatuur	16
Bijlage 1: overzicht aspecten en casestudies	18

1. CONTEXT VAN HET PROMOTIE-ONDERZOEK

Ontsluiting van geo-informatie via netwerken (Internet, mobiele netwerken) kent verschillende aspecten: data-modellering, data-opslag, data-query ('spatial query'), data-uitwisseling, netwerk-protocollen, beveiliging, software-architectuur, data-visualisatie ('Web-kartografie'), specifieke 'geo-services' (zoeken van een adres, route-planners, buffering etc).

Wat de uitwisseling van data over Internet speciaal maakt is o.a. het feit dat er niet sprake is van vaste verbindingen, maar van 'stateless' connecties. Er worden boodschappen heen en weer gezonden tussen server(s) en clients zonder dat sprake is van een 'permanente' sessie die voortduurt totdat de gebruiker uitlogt. In plaats van de bestaande GIS-software moet nieuwe applicatie-software ontworpen worden.

Internet biedt aan de andere kant ongekende mogelijkheden voor het ontsluiten en uitwisselen van geo-informatie. In plaats van kostbaar en tijdrovend data-transport op tapes en CD-ROM's zou het mogelijk moeten zijn de data te laten waar ze is: bij de bron. Een van de vragen van dit onderzoek is dan ook: hoe kan het principe 'data bij de bron' worden gerealiseerd in Web-GIS-toepassingen zonder dat concessies gedaan worden aan geboden functionaliteit, gebruiksgemak, beveiliging en 'performance'.

1.1 'Building blocks' en software-componenten

Het onderzoek 'Ontsluiting van Geo-informatie via netwerken' maakt deel uit van een groter, multi-disciplinair onderzoeksproject aan de TU Delft: het BETADE-project. BETADE staat voor 'Building blocks for Effective Telematics Application Development and Evaluation'.

Centraal in het BETADE-project staat het concept 'building block'. Onder 'building block' verstaat men een semi-onafhankelijk onderdeel van een groter geheel met een specifieke functie en duidelijk gedefinieerde interfaces naar andere onderdelen/bouwstenen. Het 'geheel' kan een administratief of logistiek proces zijn, maar ook een software-systeem. Daar waar 'building blocks' worden geïmplementeerd in software-bouwstenen spreekt men van software-componenten [Toussaint 2000].

In de literatuur over Component Based Software Development worden o.a. de volgende kenmerken van een software-component genoemd:

- de nadruk op het feit dat het om een 'runnable' object gaat (binair, of althans uitvoerbaar, dus ook een ASP-script of een HTML-pagina met JavaScript valt eronder);
- het besef dat een software-bouwsteen gedefinieerd is door de specificatie van zijn interfaces;
- qua besturingssysteem streven naar platformonafhankelijkheid, en dus;
- vervangbaarheid door een component uit een andere productfamilie (van een andere leverancier) als graadmeter of een software-object ook echt een 'building block' is.

De gedachtegang achter de componenten-benadering is niet nieuw: ook in de tijd van het gestructureerd ontwerp had men het over modulair programmeren en het specificeren van interfaces. Als criteria voor een goed ontwerp keek men naar 'minimale koppeling' tussen software-modules en 'maximale cohesie' binnen een module [Sommerville 1989]. Dit komt dicht in de buurt van de definitie van 'building blocks' als semi-onafhankelijke onderdelen die hergebruikt kunnen worden resp. vervangen kunnen worden zonder dat de rest van een programma hoeft te worden aangepast.

Ditzelfde voordeel ('hergebruik') werd genoemd toen object-georiënteerd ontwerpen als methode in opkomst was. Inmiddels zijn voor verschillende platforms omvangrijke standaard-bibliotheken beschikbaar met herbruikbare code.

Het nieuwe van de componenten-benadering is het feit dat het niet in de eerste plaats om herbruikbare code gaat, maar om herbruikbare (en vervangbare) uitvoerbare objecten. Ook de nadruk op de specificatie van de interfaces valt op (als een soort 'contracten' die vastleggen welke services een component van andere componenten verwacht en welke services de component zelf levert). Dit is een aspect dat bij object-georiënteerd ontwerpen niet expliciet de aandacht krijgt.

Het is niet verwonderlijk dat Component Based Development juist de laatste jaren opgang doet. De opmars valt samen met de opkomst van Internet en mobiele netwerken. Software-ontwikkeling werd meer dan het ontwerpen en bouwen van programma's. Naast het specificeren van de functionele eisen, moest ook nagedacht worden over de meest optimale configuratie: welke software is aan de server-kant nodig, welke aan de client-kant, hoe moet de communicatie tussen de verschillende onderdelen worden geregeld. Juist in deze nieuwe werkelijkheid komt de componenten-benadering goed tot zijn recht.

Naast een 'functionele architectuur' van de software zal ook een 'configuratie architectuur' moeten worden ontworpen. In theorie zouden individuele componenten moeten kunnen worden verplaatst naar een andere locatie in het netwerk zonder dat dit leidt tot aanpassingen in de overige componenten.

Hergebruik en vervangbaarheid van componenten is echter niet iets wat zomaar tot stand komt [o.a. Florijn 2000]. Vandaar dat in het BETADE-project apart aandacht besteed wordt aan de vraag hoe individuele bouwstenen getest kunnen worden op hun bruikbaarheid en kwaliteit als component. Oftewel: hoe garandeer je de 'composability' van een component [Hammer 2000].

1.2 Software-componenten en UML

Consequentie van de componenten-benadering is dat van te voren goed moet worden nagedacht over de architectuur van de te bouwen applicatie. De gewenste functionaliteit moet worden opgesplitst in (semi-) zelfstandige eenheden (decompositie). Als modelleer-techniek wordt vaak gebruik gemaakt van UML (Unified Modeling Language). UML is niet alleen een notatiewijze voor data- en proces-modellen, maar biedt ook de mogelijkheid op verschillende niveaus van abstractie naar data, processen en configuraties te kijken [Page-Jones 2000]. Zo is UML in alle fasen van de life cycle van software bruikbaar: van definitiestudie tot en met testen, implementatie en gebruik.

In de case studies die onderdeel zijn van dit promotie-onderzoek zal UML gebruikt worden om de functionele en technische architectuur te beschrijven van de prototypes die gebouwd zullen worden.

1.3 Geo-informatie en netwerken

Bij het ontsluiten van geo-informatie via Internet of Intranet doen zich deels dezelfde problemen voor als bij het ontsluiten van andere typen informatie:

1. hoe vindt een gebruiker de juiste informatie (catalogus-diensten en zoek-interface)
2. hoe beschrijf je als aanbieder de informatie zo dat de gebruiker een goed beeld krijgt van kwaliteit, actualiteit, eventuele verwante informatie (meta-informatie)

3. hoe zet je gegevens en meta-informatie klaar zonder allerlei handmatige handelingen (al dan niet realtime koppelen van produktie-databases)
4. hoe garandeer je randvoorwaarden als: beveiliging, snelheid waarmee de Web-pagina's worden opgebouwd, beheerbaarheid van de configuratie

Bij het ontsluiten van geo-informatie komen daar nog een aantal kwesties bij:

5. bij kaartinformatie gaat het bijna altijd om grote tot zeer grote bestanden, zodat 'performance' bij online-raadplegen nog een bottleneck is
6. GIS-pakketten kennen tot dusver hun eigen opslag-formaat, zodat het online beschikbaar stellen van bestanden van verschillende herkomst in een Web-applicatie nog veel (conversie)problemen oplevert, zo het al mogelijk is
7. standaarden voor het opvragen, uitwisselen en presenteren van geo-informatie op Internet zijn nog in ontwikkeling of zijn zo 'vers' dat de gevestigde GIS-software-leveranciers nog aarzelen met implementatie
8. de grafische presentatie (visualisatie) van geografische informatie moet aan kartografische regels voldoen; wanneer een gebruiker naar believen data uit verschillende bron kan combineren kunnen onduidelijke of misleidende digitale 'kaarten' ontstaan
9. er zullen behalve data ook 'services' beschikbaar gesteld moeten worden, zoals de mogelijkheid voor een ander referentie-stelsel / andere projectie te kiezen. Andere services zijn: het zoeken van een adres, het bepalen van de kortste route, het kiezen van een andere legenda, het plaatsen van tekst, generalisatie bij uitzoomen, incrementeel verfijnen / detaillering bij inzoomen etc.
10. bij het combineren van heterogene data uit verschillende bron is het zaak de 'betekenis' van de informatie (semantiek) eenduidig vast te leggen [Uitermark 1999].

1.4 OpenGIS en interoperabiliteit

Gelukkig zijn er de laatste paar jaar ontwikkelingen die voor een aantal van de hierboven genoemde punten een oplossing moeten bieden: er is sinds 1994 een OpenGIS Consortium waarbij een aantal grote leveranciers en kennis-instellingen is aangesloten (ESRI, Intergraph, Oracle; in Nederland ook: het Kadaster, ITC Enschede, TU Delft en Geodan). Voornaamste doel is: het zorgen voor uitwisselbaarheid van data-formaten en software-componenten van verschillende leveranciers, zodat geo-informatie op een transparante manier via netwerken kan worden ontsloten (interoperabiliteit). Het moet er met andere woorden niet toe doen waar de bestanden zijn opgeslagen of in welk pakket ze in eerste instantie zijn vervaardigd.

Om dit doel te bereiken wordt in verschillende werkgroepen geprobeerd tot eenduidige standaarden te komen voor opslag-formaat, communicatie-protocollen, grafische representatie etc. Ook worden her en der sites ingericht met daadwerkelijke toepassingen die op de standaards gebaseerd zijn: Web Mapping Testbeds [Van Oosterom 2000].

De eerste generatie Web Mapping Testbeds waren gebaseerd op het principe dat aan de server-kant een bitmap gemaakt wordt van de gevraagde kaartlagen, de bitmap wordt vervolgens als onderdeel van een HTML-pagina getoond in de browser. Elke keer dat wordt ingezoomd of attribuut-informatie wordt opgevraagd moet aan de server-kant opnieuw een bitmap worden gemaakt resp. extra informatie worden opgehaald.

Sinds 2000 wordt geëxperimenteerd met 'vector streaming'. In plaats van raster-data wordt vector-data van de server naar de client gezonden. Voordeel hiervan is dat er betere mogelijkheden zijn voor lokale interactie met de data, bijvoorbeeld het aanpassen van de legenda, het combineren met lokale bestanden, het aanpassen van kleur, lijndikte of tekstgrootte en het uitvoeren van analyses op de data (bijvoorbeeld het bepalen van de kortste route, het selecteren van objecten die binnen een bepaalde buffer vallen etc.). Inzoomen, andere lagen toevoegen of uitzetten gaat bij vector-streaming veel sneller dan bij de bitmap-benadering, het initieel ophalen kan echter langer duren, afhankelijk van de hoeveelheid data.

Onderdeel van het onderzoek is: vergelijking van de bitmap-benadering (raster-data, 'thin client') met oplossingen die van 'vector streaming' gebruik maken en een 'thick client' vereisen. (Als derde mogelijkheid is er de 'medium client', gebaseerd op formaten die eigenlijk voor grafische toepassingen ontwikkeld zijn, zoals SVG (Scalable Vector Graphics) en CGM.)

1.5 Communicatie-raamwerken

In gedistribueerde omgevingen is een belangrijke rol weggelegd voor software die de communicatie verzorgt tussen de software en data-omgevingen op de verschillende servers en clients [Bachmann 2000].

Twee bekende 'raamwerken' zijn: CORBA en DCOM. CORBA (= Common Request Broker Architecture) kan in verschillende 'smaken' worden geïmplementeerd (van 'coarse grained' tot 'fine grained'). Een voordeel van CORBA is verder dat men niet gebonden is aan een bepaalde leverancier of een bepaald platform. Dit laatste geldt niet voor Microsofts DCOM-architectuur (= Distributed Component Object Model) [o.a. Jacobsen 1998].

Uit de Java-hoek is er - naast het Enterprise Java Beans-model - als recente ontwikkeling de opmars van Jini, een verzameling API's en protocollen die ervoor zorgen dat 'clients' en 'services' elkaar op een netwerk kunnen 'vinden'.

Daar waar 'services' gedistribueerd zijn vindt ook veel data-transport plaats. Op dit vlak staan XML (als data'formaat') en SOAP (als protocol) op dit moment sterk in de belangstelling.

2. DOEL VAN HET PROMOTIE-ONDERZOEK

Hierboven is al aangegeven dat er ondanks de nieuwe technologische ontwikkelingen nog veel knelpunten zijn op weg naar open geo-informatie-netwerken. Doel van dit onderzoek is een oplossing te bieden voor een aantal van deze knelpunten. Het accent zal liggen op het ontsluiten en uitwisselen van data en de daarvoor benodigde 'services'. Hoofdvraag is: ontwerp en evalueer een generieke architectuur voor gedistribueerde geo-informatie-systemen met daarbinnen keuzemogelijkheid voor verschillende manieren van implementatie (bijvoorbeeld 'thin client' versus 'thick client'). Een belangrijke functionele eis hierbij is het principe van 'data bij de bron'. Andere eisen zijn o.a.: robuustheid (zeker bij data transacties), beheerbaarheid, uitbreidbaarheid, toegankelijkheid (gebruiksvriendelijkheid).

Relatie met OpenGIS: de binnen OpenGIS werkgroepen gespecificeerde standaarden zullen getest worden op bruikbaarheid en toepasbaarheid.

Relatie met BETADE: er moet sprake zijn van generieke oplossingen, met andere woorden: de software-bouwstenen moeten hergebruikt kunnen voor andere toepassingen waar geo-informatie wordt uitgewisseld. Ten tweede moet de vervangbaarheid worden getest: leidt het bouwen van componenten volgens interface-specificaties inderdaad tot flexibele 'geassembleerde' systemen.

3. OPZET EN FASERING

3.1 Fase 1 (eerste jaar):

In deze eerste fase ligt het accent op het inventariseren van de belangrijkste ontwikkelingen op het terrein van Web-applicaties in het algemeen en gedistribueerd GIS in het bijzonder. Achtereenvolgens worden de volgende deelvragen onderzocht:

1. Aan welke eisen moet een architectuur voor de ontsluiting en uitwisseling van geo-informatie via netwerken voldoen? Afhankelijk van de invalshoek zullen dit andere eisen zijn. Vanuit het perspectief van de gebruiker zijn snelheid data-transport, gebruiksgemak, kartografische regels, vindbaarheid (goede meta-informatie, catalog services) het belangrijkste. Vanuit het perspectief van de Webmaster en de programmeur hebben weer andere zaken prioriteit: stabiliteit en onderhoudbaarheid configuratie(s), flexibiliteit/uitbreidbaarheid software, interoperabiliteit, beveiliging en authenticatie, versiebeheer van data en software-componenten etc.
2. Welke initiatieven en voorbeelden van Internet-GIS-oplossingen zijn er in binnen- en buitenland, wat zijn de huidige knelpunten? Voor welke knelpunten wordt op dit moment een oplossing gezocht binnen het OpenGIS Consortium en andere platforms, voor welke nog niet. Evalueer de oplossingen / standaarden die er zal zijn.
3. Welke configuraties (d.w.z. hardware, besturingssysteem, server-software, middleware, browser-software) zijn mogelijk, wat zijn de voor- en nadelen?
4. Schrijf een paper voor een conferentie over de punten 1 t/m 3.
5. Ontwerp, bouw en evalueer een of twee geo-Web-applicaties gebruik makend van bestaande componenten, waar nodig aangevuld met eigen componenten. Schrijf hierover een artikel of paper voor een conferentie.

3.2 Fase 2 (tweede en derde jaar):

Omdat de ontwikkelingen op het gebied van Internet-software en protocollen en OpenGIS standaarden snel gaan is het nodig de inventarisatie-stappen 2 en 3 (zie hierboven) in het tweede en derde jaar te herhalen:

6. Herhaal de stappen 2 en 3 en bepaal welk(e) aspect(en) van gedistribueerd GIS in de case studies terug moeten komen.
7. Selecteer op basis van deze aspecten de uit te voeren case studies.
8. Voer de case studies uit.
9. Test en evalueer de gekozen oplossingen; bepaal in hoeverre de ontworpen architectuur generiek is (en de componenten herbruikbaar zijn).
10. Schrijf 4 papers voor conferenties (nog 'groene' ideeën, verslaglegging case studies)
11. Schrijf 2 artikelen voor een internationaal wetenschappelijk tijdschrift (wat gerijpte ideeën, architecturen).

3.3 Fase 3 (vierde jaar):

12. Schrijven proefschrift; hoofdvraag: hoe ziet een 'open' architectuur, waarin geo-data en geo-services 'transparant' ter beschikking staan, eruit?

4. DE CASE STUDIES (PROTOTYPES)

In het kader van het onderzoek zullen een aantal case studies worden uitgevoerd, waarvan in ieder geval twee bij de provincie Gelderland.

Bij elke case study zal globaal onderstaande werkwijze gevolgd worden:

- breng de gewenste functionaliteit en de mogelijke configuraties in beeld m.b.v. UML. Oftewel in termen van het BETADE-project: definieer de bouwstenen (organisatorische en software-matige).
- kijk welke software-bouwstenen al kant en klaar te gebruiken zijn; bijvoorbeeld: voor het ontsluiten van de metadata zouden de Catalog Services zoals geïmplementeerd in de NCGI-pilot kunnen worden gebruikt [Thewessen 2000]
- kijk welke standaarden of (concept) specificaties voorhanden zijn (OpenGIS, ISO, W3C, OMG) en bouw/assembleer prototypes gebaseerd op die specificaties
- evalueer zowel het 'assemblage'-proces als de opgeleverde architectuur en de kwaliteit van de componenten

In Bijlage 1 is een overzicht opgenomen van de geplande case studies, waarbij verticaal is aangegeven welk aspect van een gedistribueerde geo-informatie-infrastructuur vooral aan de orde komt.

Case 1: nieuw gegevensmodel voor de Topografische Kaart (Top10) van Nederland

In opdracht van de Topografische Dienst werkt de sectie GIS-technologie – samen met het ITC in Enschede en het CGI in Wageningen - aan een nieuw datamodel voor de Topografische Kaart 1:10.000 van Nederland. Dit conceptuele model moet vertaald worden naar een XMLSchema-document volgens de GML-standaard van het OpenGIS Consortium. De topografische data zelf wordt conform dit XMLSchema omgezet naar GML-bestanden. GML (Geography Markup Language) is gebaseerd op XML en zou een vervanging kunnen zijn voor de bestaande uitwisselingsformaten die door de Topografische Dienst gebruikt worden (NEN1848, ArcExport) [De Vries e.a. 2001].

Om te kunnen testen of het data-model voor de Topografische Kaart zonder informatieverlies kan worden vertaald naar GML, zal een GML-viewer gebouwd worden. De viewer moet ook de mogelijkheid geven per kaartlaag de legenda aan te brengen (kleuren, lijndikte, symbool-set, eventuele classificatie). GML is namelijk alleen bedoeld voor de opslag en uitwisseling van geometrie en attribuut-informatie ('het Digitaal Landschaps Model') en niet voor het vastleggen van classificatie en legenda ('het Digitaal Cartografisch Model'). Deze scheiding van inhoud en presentatie of 'stijl' heeft als voordeel dat een en dezelfde dataset op verschillende manieren kan worden gevisualiseerd (bijvoorbeeld soms in 2D, soms in 3D), al naar gelang de functionele wensen van de gebruiker.

Bij het OpenGis Consortium is een specificatie in de maak om ook het definiëren van de 'stijl' te standaardiseren (Styled Layer Descriptor-specificatie). In het kader van deze case study zal ook dit nieuwe 'stijl'-formaat getest worden op bruikbaarheid.

Als eerste stap zal een viewer gebouwd worden die GML omzet naar SVG-formaat (Scalable Vector Graphics), gebruikmakend van XSLT. Eenmaal in SVG-formaat kan de data bekeken worden in een standaard browser met een van de vrij beschikbare SVGViewer-plugins.

In tweede instantie zal gewerkt worden aan 'services' voor het inlezen van GML in de GIS-pakketten ArcView en MicroStation.

Door een vergelijking van de eindresultaten na conversie met de voorbeelden die als input dienden voor het GML-model kunnen conclusies worden getrokken over de

adequaatheid van het nieuwe gegevensmodel en het GML-prototype.

Doel van deze case is het testen van een aantal standaarden (OpenGIS, W3C) op het gebied van data-modellering, uitwisselingsformaten en (karto)grafische presentatie.

Case 2: het beschikbaar stellen van kadastrale gegevens via Internet

Samen met het Kadaster en de provincie Gelderland zal worden gewerkt aan een prototype voor de uitwisseling van kadastrale gegevens via Internet [Meerburg e.a. 2001].

Op dit moment krijgt de provincie Gelderland periodiek data-leveringen op draagbare media (floppies, tapes etc) van het Kadaster. Dit betreft zowel administratieve als geometrische gegevens. Verder beschikt de provincie over een modem-verbinding met het Kadaster, waarmee voor individuele percelen eigendomsgegevens (online) en kadastrale grenzen (fax) kunnen worden opgevraagd.

Deze methode van distributie heeft de volgende nadelen:

- er is veel tijd nodig voor de export, de versturing en de import/verwerking van de periodieke leveringen
- het samenstellen en doorvoeren van abonnementsaanvragen voor extra gegevens kost veel tijd en is foutgevoelig

Doel van de pilot is te komen tot een andere manier van distributie. In de pilot zullen waarschijnlijk de volgende concepten worden uitgetest:

1. data bestellen via Internet met een kaart als ingang

Onder 'kaart als ingang' moet worden verstaan dat een gebruiker op basis van een geografische selectiewijze (via een polygon of route met een bepaalde breedte) administratieve (AKR) en geometrische data (LKI) bij het Kadaster kan bestellen. De geografische selectie vindt plaats op basis van de Kadastrale Kaart die fysiek bij het Kadaster staat en 'online' bekeken kan worden door de eindgebruiker. Op basis van deze bestelling worden de gevraagde gegevens op automatische wijze verzameld ('zip and ship'). De gegevens worden via een netwerkverbinding (email en/of ftp) teruggestuurd naar de vrager. Bij de vrager zal de data op automatische wijze zodanig geconverteerd worden dat de opgehaalde gegevens bijna direct ('near realtime') bruikbaar zijn in de lokale GIS-applicaties.

Deze stap fungeert als 'opstapje' naar:

2. data bij de bron

Een stap verder dan 'kaart als ingang' is de mogelijkheid de database van het Kadaster (zowel de geometrische als de administratieve gegevens) online te raadplegen met behulp van GIS-applicaties alsof de kadastrale gegevens lokaal aanwezig zijn. De integratie met lokaal aanwezige data moet voor de eindgebruiker geheel transparant plaatsvinden, m.a.w. het doet er niet meer toe of data lokaal of op een server elders is opgeslagen.

Het is duidelijk dat deze optie vooral op het punt van gebruikers-authenticatie, beveiliging en performance de nodige eisen stelt.

Case 3: de uitwisseling van geo-informatie bij wegenprojecten

In alle fasen van een wegenproject (van concept, via grondaankoop tot en met beheer) is gedetailleerde (grootschalige) kaartinformatie nodig (eigendomsgrenzen, beheergrenzen, topografie/basiskaart). De informatie moet bovendien actueel zijn, terwijl in sommige gevallen ook eerdere versies van dezelfde kaarten bewaard moeten blijven (historische gegevens).

Het is op dit moment nog niet mogelijk op een transparante manier (zonder conversie-slagen) de verschillende soorten informatie beschikbaar te stellen buiten de software-omgeving waar de bestanden onderhouden worden. Bij wegenprojecten wordt 'van oudsher' in de ontwerpfase met CAD-systemen gewerkt; voor de 'overlay' met bijvoorbeeld bestemmingsplannen is analyse in GIS-systemen nodig. Complicerende factor is bovendien dat in verschillende fasen van een project verschillende organisaties, diensten en afdelingen bij het werkproces betrokken zijn zodat het beheer van de gegevens moeilijk centraal te regelen is.

Behalve het probleem van de 'overlay' van verschillende bestandsformaten, speelt dus ook het probleem van versiebeheer en verificatie. Onderdeel van deze pilot is te onderzoeken hoe herkomst en versie van gegevens op een adequate manier kunnen worden bewaakt. In termen van modellering: hoe kunnen temporele aspecten (historie) en versiebeheer het best worden meegenomen in ontwerp en implementatie.

Case 4: transactieverwerking via Internet

Bij de voorgaande prototypes gaat het om het transparant beschikbaar stellen van geo-informatie via netwerken voor raadpleegdoeleinden, waarbij een oplossing gevonden moet worden voor het probleem van verschillende bestands-formaten en software-omgevingen. In deze case gaat het om een onderzoek naar decentraal muteren van geografische bestanden (op bijvoorbeeld een laptop in het veld) en het verwerken van die mutaties op een 'centrale' server.

De vraag is hier hoe wijzigingen in de geometrie (door nieuwe metingen bijvoorbeeld) ingepast moeten worden in de bestaande data. Er vinden immers 'on-the-fly' was/wordt-transacties plaats. Worden alleen de gemuteerde coördinaten-paren overgestuurd naar de server of worden gehele lijn- vlak- of punt-objecten vervangen? Welke eisen stelt dit aan het (geometrisch) data-model. Zijn (unieke) object-id's nodig?

Bij mutatie-transacties via Internet moeten bovendien zaken als 'locking', commit/rollback, synchronisatie, auditing en beveiliging worden geregeld.

Net als bij case study 3 speelt versiebeheer een belangrijke rol. Onderzocht zal worden of het mogelijk is op feature (=object)-niveau kwaliteitskenmerken mee te geven (datum van de mutatie, door wie, inwinningsmethode, aanduiding van betrouwbaarheid/nauwkeurigheid)

Case: grondaankoop bij natuurontwikkelingsprojecten

Case 5: transactieverwerking via mobiele netwerken

Case 6: 'usability' van Web GIS-sites (o.a. RPD-site)

Case 7: Internet-kartografie

De cases 5 t/m 7 zullen te zijner tijd, mede op basis van de voorafgaande cases, nader worden uitgewerkt.

5. SAMENWERKING MET EXTERNE ORGANISATIES

In het kader van dit onderzoek zal met verschillende organisaties worden samengewerkt, met name met de provincie Gelderland en het ITC. Optioneel kan aansluiting gezocht worden bij interoperabiliteits-initiatieven in de praktijk: o.a. van ESRI, het NCGI (Nationaal Clearinghouse Geo-Informatie) en KLIC (Kabels en Leiding Informatie Centrum).

1. De samenwerking met de provincie Gelderland moet o.a. leiden tot een aantal Web-applicaties rond kadastrale data (één van de case studies).
De provincie Gelderland heeft tussen 1998 en 2000 meegedaan aan een aantal pilots rond Internet-ontsluiting van geo-informatie waarbij geprobeerd is volgens OpenGIS standaarden te werken:
 - het MER-project A50/A73 (samen met Rijkswaterstaat en Geodan)
 - de NCGI-pilot: 'Gedistribueerde meta-informatie' (Catalog Services gebaseerd op CORBA) (samen met het NCGI en Geodan)
 - de pilot Web Mapping gebaseerd op ArcIMS (samen Geodan)

Verder wordt bij de afdeling GEO al enige tijd geëxperimenteerd met ArcIMS (Internet-publishing vanuit ArcView): de I-kaart ("I-kaart" staat voor 'interactieve, intelligente, integrerende kaart of GIS-applicatie op Internet')

Ondanks de stappen die al gezet zijn naar een open, voor de gebruikers transparante infrastructuur voor geo-informatie, zijn er voor de naaste toekomst een aantal punten die aandacht vragen [Van Houten 1999]:

- op gegevensvlak: het inhoudelijk beheer van de geografische gegevens op de centrale server (versie management)
- op applicatie-vlak: hergebruik van software-componenten zodat applicaties sneller ontwikkeld kunnen worden
- wat betreft architectuur/configuratie: het streven naar een duidelijker splitsing van functionaliteit, gegevens en presentatie

2. Bij het ITC is recent een promotie-onderzoek gestart van Rob Lemmens dat veel raakvlakken heeft met dit onderzoek. Rob Lemmens zal zich vooral richten op het beschikbaar stellen van 'geoprocessing services' in gedistribueerde GIS-omgevingen. Uitgangspunt is de gedachte dat een GIS-gebruiker voor het oplossen van een analyse-vraagstuk (of meer in het algemeen het uitvoeren van een taak) een serie GIS-operaties nodig heeft die tesamen moeten zorgen voor het voltooien van die taak [Lemmens 2001]. Ook in het onderzoek van Lemmens speelt de interoperabiliteit van services en data – als componenten in een gedistribueerd GIS – een centrale rol.

De samenwerking zal bestaan uit inhoudelijk overleg en het uitwisselen van informatie, maar kan ook leiden tot het inrichten van een (gedistribueerde) test-omgeving voor een of meer prototypes.

De promotor van Rob Lemmens, Prof. Dr. Wolfgang Kainz, zal co-promotor zijn bij mijn promotie-onderzoek.

6. ORGANISATIE VAN HET ONDERZOEK

6.1 Opleiding en training

Elk jaar zal uit het reguliere cursus- en college-aanbod van de TU Delft een keuze gemaakt worden.

In een aantal gevallen is gespecialiseerde training buiten de TU Delft nodig. Per jaar zal hiervoor een plan en een begroting worden opgesteld. Te denken valt aan:

- opleiding Oracle9i (Java-bibliotheken, XML-streaming, inrichten Internet-architecturen)
- stage bij een buitenlands onderzoeks-instelling

6.2 Begeleiding

De begeleiding geschiedt als volgt:

- met de promotor zal 1 maal per maand overleg plaatsvinden;
- met de co-promoter zal 2 maal per jaar overleg plaatsvinden.

6.3 Ander inhoudelijk overleg

Verder wordt het volgende inhoudelijk overleg voorzien:

- wekelijkse workshops van de sectie GIS-technologie;
- elke maand overleg met de afdeling GEO van de provincie Gelderland;
- BETADE lunch-meetingen 6 maal per jaar;
- overleg met Rob Lemmens 4 maal per jaar.

7. BENODIGDE MIDDELEN

7.1 Software

Binnen het GDMC (Geo-Database Management Center) van de sectie GIS-technologie is de volgende software beschikbaar:

- (OO)DBMS-software van Oracle, Computer Associates (Ingres, Jasmine) en Informix
- GIS en CAD-software van ESRI en Bentley
- Internet-GIS-software van PGS, ESRI en Bentley

Voor dit promotie-onderzoek is verder in ieder geval de volgende software nodig:

- Rational Rose (UML-tool)
- XMLSpy of TurboXML (XML-validator plus XSLT-parser)
- FME (voor het gemakkelijk inlezen van testbestanden in de verschillende database-omgevingen)
- EndNote of andere bibliografische software
- software die tijdens het onderzoek relevant blijkt te zijn als hulpmiddel of als te testen produkt

7.2 Kosten congresbezoek

Naar verwachting zullen er per jaar in ieder geval twee buitenlandse congressen worden bezocht.

7.3 Kosten bezoek aan binnenlandse en buitenlandse instanties en onderzoeksinstellingen

Verder worden bezoeken aan instanties en onderzoeksinstellingen in Nederland en het buitenland voorzien.

7.4 Kosten gespecialiseerde opleiding en training

Ook voor kosten gespecialiseerde opleiding en training dient een bedrag gereserveerd te worden.

De kosten worden in principe door het DIOC BETADE vergoed.

Resumé

<i>Titel onderzoek</i>	<i>ONTSLUITING VAN GEO-INFORMATIE VIA NETWERKEN</i>
<i>Promovendus</i>	<i>Marian de Vries</i>
<i>Promotor</i>	<i>Prof.Dr.Ir. Peter van Oosterom</i>
<i>Co-promotor</i>	<i>Prof.Dr. Wolfgang Kainz</i>
<i>Doel onderzoek</i>	<i>Ontwerp en evalueer een generieke architectuur voor gedistribueerde geo-informatie-systemen waarbij 'transparante' uitwisseling van data en services tussen heterogene omgevingen mogelijk is. Vergelijk een aantal mogelijke technische implementaties. Formuleer functionele eisen en stel evaluatie-criteria op.</i>
<i>Financiering</i>	<i>DIOC BETADE TU Delft Provincie Gelderland Faculteit CiTG, Afdeling Geodesie, sectie GIS technologie TU Delft</i>
<i>Periode</i>	<i>November 2000 – november 2004</i>

Literatuur

Bachman, F., L. Bass, C. Buhman e.a. (2000), *Volume II: Technical Concepts of Component-Based Software Engineering*, TECHNICAL REPORT CMU/SEI-2000 TR-008 ESC-TR-2000-007, May 2000

Online: <http://www.sei.cmu.edu/pub/documents/00.reports/pdf/00tr008.pdf>

Conallen, J. (1999), *Building WEB applications with UML (The Addison-Wesley Object Technology Series)*, Addison-Wesley, 1999

Van Houten, W. en M. van der Ven (1999), *Gelderland geeft ruimte aan informatie: Ruimtelijke Informatievoorziening in de provincie Gelderland*, uitgave ter gelegenheid van het Geodesia-congres 1999, oktober 1999

Florijn, G. (2000), "Component Based Development: hype of hoop?", in: Het Ingenieursblad, 5/2000

Geography Markup Language (GML), Version 2.0, OGC Recommendation Paper, 20 February 2001

Hammer, D.K. and M.R.V. Chaudron (2000), "Towards Component-Based Architecting for Resource Constraint Systems", paper 4th International Software Architecture Workshop, Limerick, June 4-5, 2000

Jacobsen, H.-A. and A. Voisard (1998), "CORBA-Based Interoperable Geographic Information Systems", paper 2nd IASTED European Parallel and Distributed Systems Conference, Vienna, 1998

Lemmens, R., (2001), *Distributed interoperable processing in GIS: Towards Plug-and-Play GIS components*, PhD research proposal, January 2001

Meerburg E., M. de Vries, J. Stufken, B. Maessen (2001), *Functionele analyse distributie Kadaster – provincie Gelderland (concept)*, maart 2001

Van Oosterom, P.J.M. en M.W. Schram (2000), 'OpenGIS: wat is het en wat betekent het voor u?', in: *Geodesia*, 2000-4

Orfali, R. and D. Harkey (1998), *Client/Server Programming with Java and CORBA*, John Wiley & Sons, 1998

Page-Jones, M. (2000), *Fundamentals of Object-Oriented Design in UML*, Addison-Wesley, 2000

Sommerville, I. (1989), *Software Engineering*, Addison-Wesley, 1989

Thewessen, T. e.a. (2000), *Evaluatie NCGI Pilots: Gedistribueerde Metadata Architectuur*, maart 2000

Toussaint, P. e.a. (2000), *BETADE position paper (work in progress)*, februari 2000

Uitermark, H.T., P.J.M. van Oosterom, N.J.I. Mars, M. Molenaar (1999), *Ontology-Based Geographic Data Set Integration*, paper Workshop on Spatio-Temporal Data Base Management Systems (STDBM'99), Edinburgh, UK

De Vries, M., T. Tijssen, J. Stoter, W. Quak, P. van Oosterom (2001), *Topographic data, object orientation and GML*, paper EOGeo 2001

XSL Transformations (XSLT) Version 1.0. W3C Recommendation, november 1999

Bijlage 1: overzicht aspecten en case studies

	Case 1 Topographic data	Case 2 Cadastral data	Case 3 Road projects	Case 4 Distributed update	Case 5 Wireless	Case 6 User interface	Case 7 Multi- source cartography
Data modeling	X	X	X				
Data storage	X	X					
Data query	X	X	X			X	
Data transfer		X	X	X	X		
Heterogeneous data		X	X			X	X
Data management		X	X	X			
Metadata Catalog Services			X			X	X
Distributed geo-services		X	X	X	X		X
Visualization	X				X	X	X
Semantics		X	X			X	X
Topology	X	X	X	X			
Temporal aspects	X	X	X	X			
Update transactions				X			
Update 'propagation'				X			
Distributed architecture		X	X	X	X		
Heterogeneous environments		X	X	X	X		
Communication framework		X	X	X	X		
Security		X	X	X	X		
User authentication		X	X	X	X		
e-commerce		X				X	
OpenGIS standards	X	X	X	X			X
BETADE concepts and methodology		X	X	X	X	X	X

Eerder verschenen rapporten in deze reeks:

- GIST Report No. 1, Oosterom, P.J. van, Research issues in integrated querying of geometric and thematic cadastral information (1), Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 29 p.p.
- GIST Report No. 2, Stoter, J.E., Considerations for a 3D Cadastre, Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 30.p.
- GIST Report No. 3, Fendel, E.M. en A.B. Smits (eds.), Java GIS Seminar, Opening GDMC, Delft, 15 November 2000, Delft University of Technology, GIST. No. 3, 25 p.p.
- GIST Report No. 4, Oosterom, P.J.M. van, Research issues in integrated querying of geometric and thematic cadastral information (2), Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 29 p.p.
- GIST Report No. 5, Oosterom, P.J.M. van, C.W. Quak, J.E. Stoter, T.P.M. Tijssen en M.E. de Vries, Objectgerichtheid TOP10vector: Achtergrond en commentaar op de gebruikersspecificaties en het conceptuele gegevensmodel, Rapport aan Topografische Dienst Nederland, E.M. Fendel (eds.), Delft University of Technology, Delft 2000, 18 p.p.
- GIST Report No. 6, Quak, C.W., An implementation of a classification algorithm for houses, Rapport aan Concernstaf Kadaster, Delft 2001, 13.p.
- GIST Report No. 7, Tijssen, T.P.M., C.W. Quak and P.J.M. van Oosterom, Spatial DBMS testing with data from the Cadastre and TNO NITG, Delft 2001, 119 p.
- GIST Report No. 8, Vries, M.E. de en E. Verbree, Internet GIS met ArcIMS, Delft 2001, 38 p.
- GIST Report No. 9, Vries, M.E. de, T.P.M. Tijssen, J.E. Stoter, C.W. Quak and P.J.M. van Oosterom, The GML prototype of the new TOP10vector object model, Report for the Topographic Service, Delft 2001, 132 p.
- GIST Report No. 10, Stoter, J.E., Nauwkeurig bepalen van grondverzet op basis van CAD ontgravingsprofielen en GIS, een haalbaarheidsstudie, Rapport aan de Bouwdienst van Rijkswaterstaat, Delft, 2001, 23 p.
- GIST Report No. 11, Geo DBMS, De basis van GIS-toepassingen, KvAG/AGGN Themamiddag, 14 november 2001, J. Flim (eds.), Delft 2001, 37 p.
- GIST Report No. 12, Vries, M.E. de, T.P.M. Tijssen, J.E. Stoter, C.W. Quak and P.J.M. van Oosterom, The second GML prototype of the new TOP10NL object model, Delft, 2002.
- GIST Report No. 13, Vries, M.E. de, T.P.M. Tijssen en P.J.M. van Oosterom, Comparing the storage of Shell data in Oracle spatial and in Oracle/ArcSDE compressed binary format, Delft, 2002, .72 p. (Confidential)
- GIST Report No. 14, Stoter, J.E., 3D Cadastre, Progress Report, Report to Concernstaf Kadaster, Delft 2002, 16 p.
- GIST Report No. 15, Zlatanova, S., Research Project on the Usability of Oracle Spatial within the RWS Organisation, Detailed Project Plan (MD-NR. 3215), Report to Meetkundige Dienst – Rijkswaterstaat, Delft 2002, 13 p.
- GIST Report No. 16, Verbree, E., Driedimensionale Topografische Terreinmodellering op basis van Tetraëder Netwerken: Top10-3D, Report aan Topografische Dienst Nederland, Delft 2002, 15 p.
- GIST Report No. 17, Zlatanova, S. Augmented Reality Technology, Report to SURFnet bv, Delft 2002, 72 p